



1959

# An Empirical Avaluation of the Theory of Manifest Structure Analysis

Lucille Armstrong Foster  
*Loyola University Chicago*

## Recommended Citation

Foster, Lucille Armstrong, "An Empirical Avaluation of the Theory of Manifest Structure Analysis" (1959). *Dissertations*. Paper 585.  
[http://ecommons.luc.edu/luc\\_diss/585](http://ecommons.luc.edu/luc_diss/585)

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Dissertations by an authorized administrator of Loyola eCommons. For more information, please contact [ecommons@luc.edu](mailto:ecommons@luc.edu).



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License](https://creativecommons.org/licenses/by-nc-nd/3.0/).  
Copyright © 1959 Lucille Armstrong Foster

AN EMPIRICAL EVALUATION  
OF THE THEORY OF MANIFEST  
STRUCTURE ANALYSIS

by

Lucille Armstrong Foster

A Dissertation Submitted to the Faculty of the Graduate School  
of Loyola University in Partial Fulfillment of  
the Requirements for the Degree of  
Doctor of Education,

January

1959

## LIFE

Lucille Armstrong Foster was born in Chicago, Illinois, October 29, 1908 and was named Mary Lucille Armstrong.

She was graduated from the D.S. Wentworth Elementary School in 1922, the Parker High School in 1926 and Crane Junior College in 1928, all public schools in Chicago. The University of Illinois granted her a Bachelor of Science degree in 1930. The same year she started her teaching career in the Manito Illinois Consolidated High School where she taught Latin, mathematics, girls' gymnasium, and coached the mixed chorus, dramatics, and the girls' and boys' tennis teams.

In the next ten year period she substituted in both the Chicago public and parochial high schools and junior high schools. She was a regularly assigned Latin teacher at Crane Evening School for five of the ten years. She also was an assigned junior high school teacher, an assigned elementary school teacher and an assigned high school teacher. She took college work at Crane Junior College, Chicago Normal College, (now Chicago Teachers College), Chicago Academy of Fine Arts, De Paul University, and University of Wyoming. She also received her Master of Science degree at De Paul University in 1940. During this period she also married George E. Foster.

From 1940 to the present time she has taught mathematics and Latin at Englewood High School and has helped to raise eleven children who were without parents or homes. She also attended summer and evening courses at the University of Mexico, the University of Illinois, Knox College, University of Chicago, Northwestern University, Illinois Institute of Technology, Roosevelt University, the University of Wisconsin, Chicago Teachers College, San Francisco State College, and Loyola University.

## ACKNOWLEDGMENTS

The writer wishes to express her appreciation for the help, advice and encouragement she received from a large group of people. Her advisor, Dr. Samuel Mayo of Loyola, was always available for conferences, for reading drafts of the dissertation, for suggesting other points of view or other sources of materials. Other members of her committee--Drs. Valenti, Rimoldi, Englehart, and O'Mara also offered constructive criticism, encouragement, and advice. Dr. Kiniery and his staff of the graduate school office at Loyola also gave suggestions.

At the Cooperative Research Project Dr. David Itkin and his statistician, Mr. Brauer, took great pains to make available the data and to explain certain aspects of it. The staff also was helpful in explaining the aspects of the work of which each one was a specialist.

The friends of the writer were gracious in offering themselves as "guinea pigs" to see if untrained people could carry out most of the work of mounting, sorting, cataloging, etc. In this group the writer should mention Mrs. Armstrong, Brown, Floyd, and Hitchcock and also Misses Sekera, Brown and Hitchcock.

In checking of the mathematical calculations Misses Sekera and Brown and Mr. Foster were most helpful. Also Dr. Goldman and Professor Blaisdell of Roosevelt University gave suggestions about statistical matters.

Dr. Frank du Mas and Derek Hasse, a graduate student who was writing a thesis under his direction at Montana University helped the writer to understand matters which were not entirely clear in the book, Manifest Structure Analysis.

For the photography, drawing, and typing an expression of appreciation is due to Mr. George Foster (the writer's husband) and to his staff especially his stenographer, Mrs. Mary Traficano.

## TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION	
A. Statements of the purpose, value, and usefulness of the study . . . . .	1
B. A review of the related literature . . . . .	2
II. PRESENTATION OF THE THEORY OF MANIFEST STRUCTURE ANALYSIS . . . . .	20
III. MATERIALS AND SPECIAL TOOLS USED	
A. A description of the source of data--The Cooperative Research Project . . . . .	30
B. A description of the writer's scaling board . . . . .	31
C. A description of the chart for ascertaining degree of communality . . . . .	33
IV. THE APPLICATION OF THE THEORY AND THE RESULTS	
A. A step by step account of the application of the theory . . . . .	35
B. The use of the special tools . . . . .	40
C. The results . . . . .	44
V. CRITICAL EVALUATION	
A. A summary of	
1. Strengths and advantages . . . . .	56
2. Weaknesses and disadvantages . . . . .	56
B. Suggestions for further research . . . . .	65
BIBLIOGRAPHY . . . . .	67
APPENDICES . . . . .	69

## LIST OF TABLES

Table		Page
1	DISTRIBUTION OF 264 CASES . . . . .	36
2	DISTRIBUTION OF 264 CASES IN THIRTEEN GROUPS . .	37
3	DISTRIBUTION OF 264 CASES IN FOUR INTERVALS . . .	38
4	GRADE PLACEMENT . . . . .	39



## LIST OF FIGURES

Figure		Page
1	A COMPARISON OF THE GUTTMAN AND LAZARSFELD MODELS. .	15
2	LAZARSFELD'S TRACE LINES . . . . .	16
3	EXAMPLES OF SEGMENTAL CATESCALES . . . . .	25
4	EXAMPLES OF INTENSIVE CATESCALES . . . . .	26
5	EXAMPLE OF A CLUSTERY CATESCALE . . . . .	27
6	EXAMPLES OF MAJOR AND MINOR ENDS OF ASSOCIATION SURFACES . . . . .	49

## CHAPTER I

### INTRODUCTION

The purpose of this study is to evaluate the effectiveness of "Manifest Structure Analysis" (4) as a method of scaling which can be used on unusual assortments of data in such a way as to eliminate the extraneous information and to cut down on the amount of data necessary for success in prediction. The present study has used success in reading in the case of mentally handicapped children as a manifest variable. It has tried to show what data are important in predicting success in reading for the mentally handicapped before they have reached the mental age when they can reasonably be expected to show their ability to read.

As a general rule more research could be carried on and should be carried on by ordinary teachers. This has not been done because the people who are close to the problems and do search for solutions do not have the time nor experience to use the tools of research. du Mas (4) has proposed a kind of solution which he says does not require an extensive knowledge of mathematics and statistics. He has proposed what seems to him a simple way of discovering the salient features or the facts pertinent to the solution of a problem which involves prediction. He believes that people can learn to use his method pragmatically without the necessity of understanding the theory on which it is based.

Because Manifest Structure Analysis is purported to be a theory and method of scaling, it is advisable to cite what some others have done. Among the scaling theorists the writer has chosen to consider are Thurstone, Likert, Remmers, Guttman, Coombs, and Lazarsfeld.

Scale theories are integral parts of the behavioral sciences because of the unique problems of measurements in these fields. Sociology, psychology, and education have aspired to be classified as sciences. The qualifying adjective, behavioral, has usually been added to distinguish them from the physical sciences. The physical sciences--chemistry, physics, astronomy hold the top position, being nearer the ideal of an exact science. They use mathematical models and have had greater success in prediction than have the second ranking sciences, the biological sciences. Lord Kelvin believed that mathematics was needed to describe a phenomenon, and if the description could not be made in measured quantities, the subject or field of study could not be called a science. Mathematics has been used in various ways--sometimes in the form of graphs to present statistical results, sometimes in providing models, sometimes in giving aid in analyzing data, and sometimes in supplying a language about which there is little disagreement.

Since the behavioral sciences in the past have had a desire to emulate the physical sciences in the matter of success in prediction, and since the physical sciences have used mathematics

as a tool for achieving success in prediction, it seems logical to the writer that the behavioral sciences would need to use mathematics. Thorndike insisted that whatever exists in some measurable quantity. However, in order to have absolute measures, there must be established an absolute zero, there must be equality of units, and the property must be additive.

In the field of education, the earliest fairly successful effort at "quantifying" was the measurement of intelligence. The I.Q. has a scale which does not have an absolute zero, nor equal units, nor properties which are additive. In certain cases I.Q.s of pupils in a class are added and an average for the class is found. This has been done with the assumption that the units are equal and can be added. No one will actually say that the units between a 60 I.Q. and an 80 I.Q. represent the same intellectual distance as do the units between a 160 I.Q. and a 180 I.Q. In spite of its short comings, the scale of the I.Q. has been successful in giving a numerical description which is stable, reliable, discriminating, and useful. So far as the usefulness to the schools, colleges, and the prospective employers of students nothing comparable has been done in the measurement of attitudes, personality, or character.

Research in the field of measuring human characteristics has always been given more attention and financial assistance in war times. The urgency or the pressure of the need has produced techniques for prediction that were not mere guess work or acci-

dent but were not explained to the satisfaction of mathematicians. Measurement of the physical attributes of people--such as height, weight, finger prints has been achieved with a very satisfactory degree of consistency. Intelligence,<sup>1</sup> achievement, manual dexterity, and special aptitudes have been tested with somewhat satisfactory degrees of effectiveness. But during the two World Wars psychologists were asked to do more. They were able to tell which candidates for airplane pilots training were intelligent enough, manually dextrous enough, quick enough in their reactions. However, they were also asked to discover which ones would "crack up" under the strain of flying mission after mission and which ones would unwittingly give away valuable information if captured by the enemy. The psychologists were asked to measure morale, public opinion, and attitudes.

When Thurstone in 1929 was working on the construction of an attitude scale (12, p. 214), he assumed that a large number of questions would be better than just a few. He sought to put these questions along a continuum so that one extreme would be favorable, the middle indifferent, the other extreme unfavorable. Thurstone arranged his scale so that it would have eleven gradations or steps. He collected a number of statements about a subject (Religion) which he considered suitable for scaling. He

<sup>1</sup>Army Alpha & Army Beta published in 1915 were the earliest group intelligence tests (2, p. 5). The Army General Classification test 1940, 1941 revision (2, p. 547).

presented these statements to a number of people whose opinion he valued as appropriate, asking them to put the statements in one of eleven categories or gradations. He made frequency distributions to show the judgments for each item. Ogives were constructed; scale values at the 25th, 50th, and 75th percentiles were determined. The 50th percentile was deemed the scale value of the item and indicated its position along the measurement continuum. The 75th minus the 25th percentile yielded the Q value or the variability of the judgment or the ambiguity of the item. After that he selected 20 or 25 items which were low in ambiguity and were equally spaced along the continuum. The person who took the test would indicate which of the items he agreed with. His score would be the scale value of the median item agreed with.

The weaknesses of the Thurstone method were as follows:

1. The quality of the questions were dependent upon the experience of the author with various socio-economic classes and with various I.Q. levels;
2. Experts had to be involved to place these questions in their position along the continuum;
3. There was no way of equalizing or equating the units;
4. There was no absolute zero;
5. The units were not additive;
6. The method was not easily transferred to other fields of research;
7. The method was not applicable to data which have heterogeneous content.

The value of the Thurstone method was that he did present a solution to the problem of quantifying qualitative material. He stimulated social scientists to think of other ways of scaling.

His method was less subjective than any scaling which had been attempted before.<sup>2</sup>

In 1932 Likert (12, p. 218), like Thurstone, began his scaling of attitudes by devising statements. He did not use judges but rather presented his statement to his subjects and asked them to express their opinion as to these statements one by one. For each statement he would check one of these five opinions--strongly agree, agree, undecided, disagree, strongly disagree. Strongly agree might be weighted as five, agree as four, undecided as three, etc. The subject's score would be the sum of the weights on all the questions. Likert had a number of subjects and took the lowest and the highest 10 per cent. For these groups responses to each item were compared with the total score. Items which best differentiated between the two groups were used to compose the revised scale.

Likert's method of scaling had several of the same weaknesses as had Thurstone's, namely: 1. The suitability of the questions for all socio-economic classes were dependent upon the experience of the author; 2. The units or measures of the differences in the intensity of feeling say between strongly disagree and disagree or between disagree and undecided were not equal; 3. There was no absolute zero; 4. The units were not additive. However,

<sup>2</sup>Bogardus' Social Distance Scale 1925 preceded Thurstone's scales of 1929. (1)

Likert did get away from using the opinions of experts. He did present a system which could readily be used in other fields of research. His method was a way of bringing a wide variety of material into a situation where latent relationships might be shown to exist. In this respect his data often revealed unexpected relationships. When a score was arrived at for each person, it was not necessarily an unvarying characteristic. For example, a person who scored 25 today might not score 25 tomorrow because he might have changed his opinion on certain aspects or questions. However, certain answers might always be the same. This stability or lack of it would be measurable.

Remmers (14) devised a generalized scale. His rationale was to create a single scale which might be standardized and could be used to measure attitudes toward any specific object or phenomenon which is a sub-specie of the general class for which the scale was devised. The purported advantage was that once a set of items (statements) had been scaled, it could be used to measure attitudes of any sub-specie.

Silance<sup>3</sup> (12, p. 217) who extended the Remmers' idea had 150 college students sort 150 statements such as, "I hate Y subject," "I like to study Y subject." These were to be sorted for school subjects in general not for any particular subject. These 150

<sup>3</sup>Webb gives a very thorough description of Silance's work in developing Remmers' scale as a preliminary to his own work. (18)



students were to do what Thurstone's judges did. The questions were then given to subject matter classes, Botany, Chemistry, physics, etc. and "Y" was replaced by Botany or Chemistry or Physics. For each of the different subjects, there might be a different Q (quartile deviation) value. The scale was tested on sub-species all in the field of science. The Chi square test was also used to test the generality of an item. When the different groups, each group with a different sub-specie judged an item, the tabulations formed contingency tables and the Chi squares were computed. By referring to Pearson's tables, a P value was obtained for each item and this expressed the probability by which the sortings differed by chance. Since scale and Q values were dependent on the frequency distribution of judgments, the P value would then be an index of the significance of the difference between the scale and Q value for different sub-species.

To make the Chi square test the items had to be judged by the Thurstone method. The number of categories was reduced from 11 to 5 and each category was assigned a verbal description and a number. As in the Likert scale one was very unfavorable, 2. unfavorable, 3. neutral or indifferent, 4. favorable and 5. very favorable. An IBM card could be used with five choices.

Webb (18) had taken over the problem of scaling from the point where Silance had left it--namely the 150 items were arranged. Webb kept sixty of the Silance items intact, changed twelve and added fifty eight. He tried to be very sure that he

had equal units along a continuum. His scales were tried out on science classes at North Carolina University. His results showed the P values below the 10 per cent or even 25 per cent level of confidence were fewer than might be expected by chance on the basis of sampling theory. If there were no items with the same Q value, then the P values would be used to determine which item should be used. Webb finally arrived at a 45 item test which when correlated with two self rating scales on interest and liking for the subject come out to be .89 for one and .90 for the second self rating scale.

Remmers' scale was a study and application of semantics. Remmers, Silance and Webb were all trying to find expressions of graded feeling toward or away from something. Webb's results seem to indicate that the words or expressions used in the Remmers' scale had about the same connotation and gradations for the college students taking the science courses. Although it was not tried on subjects in other fields, it might be said to work for all college students in all subject matter fields taken in schools. At this point there is a problem: Would this scale work if it were tried on the unselected so called "man on the street," with the name of a political party substituted for Y subject? The Thurstone and Likert scales have been used in experiments which use people other than college students as subjects. Remmers has broadened his opinion polls to include high school pupils but not much has been done by him with adults who

are not college students.

Remmers, Silance and Webb used students as Thurstone had used experts as judges. This improved the semantics or made the scale closer to the thinking of a more numerous group, the ordinary college student. They wrestled with the problem of equal units and solved it to some extent. They tried with some degree of success to meet the problem of the dependency of the effectiveness of the statements on the calibre of the intellect of the author. They did not achieve absolute zero nor additive units.

Coombs (2, 3) wrote about scale theory from another point of view based upon Stevens' scale types (16). He described the scales and then suggested how the data could be fitted to the scales or how certain scales were of use in connection with certain data. For example, he said the nominal scale consisted of substituting numerals for real objects. He suggested that this be used for occupational families or psychiatric classifications. The psychological processes of perception were often representative of the measurement on a nominal scale.

Next in the order of complexity Coombs placed the partially ordered scale. In this it is seen that some members are more than just different from another class. For example, A is greater than B because A has a better education and more money than B.

Next higher was the Ordinal scale. Sometimes B and D cannot be compared because B has more money than D but less education.

In order to deal with this situation \$1000.00 of income might be equated to one year of college education. With this adjustment B and D could be placed on the Ordinal scale.

Next in complexity was the Ordered Metric scale. Here the distance between classes was involved but the units are not claimed to be exactly equal. The 150 statements of Silance and the 45 statements of Webb and the Remmers scale would be in this classification. The various I.Q. tests would also fall in this group.

A step above the Ordered Metric was the Interval scale. It was characterized by the fact that the data contained information on just how large the intervals between all the stimuli were. There was a common, constant unit of measurement. Numbers might be associated with the positions of the stimuli and arithmetic might be performed on the differences between the numbers and numbers might be added to the scale scores. The scales might be multiplied by any given number and the relationship between the numbers would be preserved.

Higher than the Interval scale was the Ratio scale. This scale differed from the Interval scale only in the matter of the zero. The Interval scale had an arbitrary zero while the ratio scale had an absolute zero. The centigrade thermometer would be an example of an Interval scale for zero is arbitrarily fixed.

Whenever the Ordered Metric or any of the simpler scales below it are used, the measurement or rating is said to be done

by the scale method or according to scaling theory. There are two points of view which can be followed, the person builds a scaling method to fit his data or he alters his data to fit the scale. He calls "error" all which does not fit the scale.

Because of the special problems connected with getting information by observation or inference, Coombs has used two names to describe the things psychologists are observing, measuring, or trying to describe. The two levels of description are genotypic in which the thing measured is inferred, hypothetical, or latent and phenotypic when the thing measured is observed or manifest. His theory of data is that there must be a definition of information contained in an observation on the phenotypic level. There must also be a definition of the relationship of the phenotypic and genotypic. This provides the basis for making genotypic inferences from the observations.

Coombs believes that the method of collecting data determines what information they contain but the method of analysis defines the material. The method of analysis may permit the discovery of the properties of the information or it may define the properties. In the latter case the experimenter wishes only to know about interrelations.

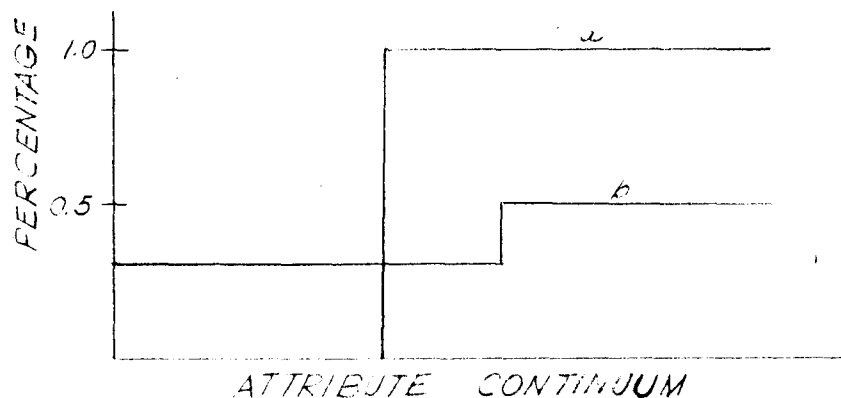
Of all the scale theorists Coombs has seemed to give no more attention and space to his own contributions--the genotypic and phenotypic theory of data, his unfolding technique, his various matrices than he has to the contributions of others. He seems

objectively to state the limitations beyond which his techniques are not useful. He presents the theories of others in the most favorable manner but mentions their limitations. He seems to say in effect that there are many tools, and that one should select the tool which best meets his needs. He started with the simple familiar mathematical scales and showed how each was particularly suited to certain data. He has taken the point of view and has had courage to write that the ratio scale--which has all the assets of equal units, absolute zero, additive units and which has been highly regarded by physical scientists and mathematicians--is not of much use to the social scientists. He reminds scientists that in fitting the data to a scale, certain data not fitting the pattern must be discarded; these data are often valuable and would contribute useful information. Coombs believes that if the data will fit a conventional or already discovered technique, this is fortunate but if they do not, a pattern or technique should be built or adapted to them, not they be adapted to a pattern or technique. Because of this belief, Coombs' Chapter (3) is a good review of the work of scale theorists and a kind of exhortation for people with unusual data to build their own scales, theories, or patterns, for them not slavishly to follow the physical scientists but boldly to discover or invent a new theory or pattern.

Guttman's (12, p. 220, 10) scale methods were put into practice during World War II. His purpose was to determine by empiri-

cal means whether an attitude was scalable. If a number of soldiers (or in some cases civilians) responded in a consistent way, then the attitude would be deemed scalable. He had a mechanical device called a scalogram by which he ascertained the degree of consistency of an item. His criterion of consistency was that endorsement of a given item was to be accompanied by the endorsements of all other items less extreme and the rejection of all items that were more extreme. When a set of items had sufficient consistency, the scale was called unidimensional and could be expected to yield reliable measures. The offending items would be thrown out until the scale met the criterion of consistency. Festinger (12, p. 221) criticized this method by saying that often the items become mere repetitions of the same idea in different words. This same criticism would apply equally well to the Thurstone, Likert, Remmers, and the du Mas methods of scaling. The Guttman method of eliminating extraneous materials as non-scalable is based on logic, but the du Mas method is not clearly defined. The du Mas method in its three types of catescales (patterns) really gets a weight or value for almost every item so that rarely is any item discarded as entirely worthless or inappropriate. It is true that some items which do not fit any of the patterns are not used. However, du Mas does seek to avoid the weakness of Guttman's method--namely that the statements are the same idea repeated in a variety of different ways.

Lazarsfeld (13) has thought of his data as multidimensional, as had du Mas but graphically he could depict only three dimensions. He has tried to let his data lead him to discover relationships. Coombs (3) felt that Guttman's scale was a special case of the Lazarsfeld Latent Structure Model. Lazarsfeld's theories are really metatheories--master theories--theories large



a: GUTTMAN TYPE  
b: LAZARSFELD TYPE

FIGURE 1. A COMPARISON OF THE GUTTMAN AND LAZARSFELD MODELS.

enough to include all lesser theories as special cases. Guttman's mental test theory was a special case of Lazarsfeld's Latent Structure Analysis theory. Lazarsfeld assumed that on a continuum there is a point on one side of which will be the people who agree and the other side will be those who disagree. The Guttman model demanded that the  $P_j$  (people endorsing the position or item) be equal to zero or one. See Figure 1 (line a). Lazarsfeld did not have to have such consistent items or reliable items. Lazarsfeld's solution provided a set of two latent classes on the geno-



typic level (Figure 1, line b). When the data satisfy the conditions for a simply ordered scale, Lazarsfeld's system reduces to Guttman's. Guttman's system is very like Coombs' Parallelogram technique. Lazarsfeld regarded the underlying attribute not as having discrete steps or classes but as being a continuous gradation. He intended his theory to apply to non-monotone<sup>4</sup> items

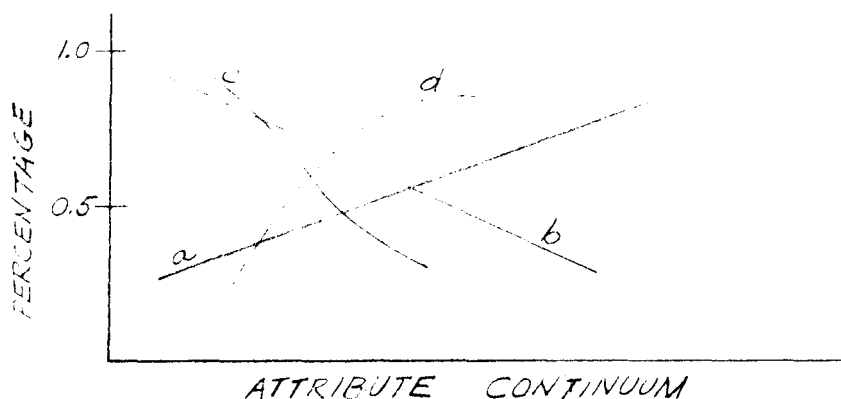


FIGURE 2. LAZARSFELD'S TRACE LINES

but it appears to apply equally well to monotone items and to a combination of them in the same test or questionnaire. In Figure 2 a, b, c are trace lines for monotone items and d for a non-monotone item. An infinite variety of trace lines may be assumed.

<sup>4</sup>Coombs used the expression 'monotone item' to refer to a question or stimulus which could be expected to discriminate as would an arithmetic problem between the careless and the ignorant on the one hand and the careful and understanding student on the other. An example of a non-monotone item would be the statement "We should make the loan to Britain if we are sure they will pay the loan to Britain with no conditions attached and by those who did not want to make the loan at all. The item discriminated between the moderates and the extremes but the extremes were lumped together in one category.

Each item has its own trace line.

According to Coombs Lazarsfeld's theory is also a generalization of Thurstone's Multiple Factor Analysis. Lazarsfeld has undoubtedly created metatheories into which the theories of others fit like a jig saw puzzle. Others except for Coombs have viewed the field in a narrow way and created devices and theories which have served their purposes. Theorists whose theories were developed because they needed them for their particular task or job were practical men. Guttman had a task to do in connection with World War II he had to have a pattern useful for prediction which would fit the data he already had or could get quickly. He could not take much time to think about how effective this pattern or theory would be on other material. His theory would be a tool built for a specific not a general purpose. He would be a practical man, a person who contributed to the field of applied psychology. Lazarsfeld's theories were too philosophical, too intricate mathematically, and too broadly general to pin point or suggest where or how they were to be used.

There has been a change in viewpoint since 1937 when Zubin according to Hasse (11) differed from his predecessors on the matter of emphasis on items passed. He maintained that a knowledge of the response configuration was more significant than items passed. The advocates of configural or pattern analysis have effected this change in viewpoint. Guttman (4, p. 7 and 11) contributed the first configural scale. Loevinger (4, p. 6) contributed formulae and criteria for the evaluation of test and item

homogeneity. All her scale models formed triangular matrices. Her methods were more rigid than Guttman's, for she was interested only in plus, or correct, answers. Lazarsfeld thought that Guttman's theories were too rigid and restrictive.

du Mas in his book, Manifest Structure Analysis (4) stated that he was determined to find a way to organize and evaluate the wealth of data collected by clinicians. He felt that Lazarsfeld's theories came the nearest to filling his purposes but they needed to be analyzed further, the complexity of the mathematics simplified, and more empirical research done in Latent Structure Analysis. He felt that Lazarsfeld's method required that items and stimulus materials, or the responses to them, all belong to a certain particular universe, or domain, and that the items all exhibit a phenomenal order.

The method of du Mas is little concerned with what the data are like, how they were gathered, or how much or how little there is of them. The method does not care what units are used to express the manifest variable, or criterion. It does suggest that better results can be expected if the cases used to build the patterns are distributed in a rectangular fashion not in a normal curve distribution. It also suggests that the criterion be selected thoughtfully and that it be readily measurable. The units used as measures of the criterion will be the same units found in the prediction. Once the criterion or manifest variable has been chosen, du Mas offers three patterns which the research

person can try to build from his data. These patterns are the ways of selecting the pertinent from the extraneous data. A more elaborate description will follow in Chapter Two.

## CHAPTER II

### PRESENTATION OF THE THEORY OF MANIFEST STRUCTURE ANALYSIS

This chapter will be concerned with a more detailed exposition of Manifest Structure Analysis, which was introduced in the previous chapter. In his introduction, du Mas says of his book (4, p. 1) "This treatise attempts to define a set of operations in which it is possible to utilize categorical or enumerative data in a quantitative scale." He also says that of the two major aims of quantitative science, prediction and measurement, measurement is basic. du Mas has attempted to deal with the problem of scaling data which neither exhibits phenomenal order nor seems to belong to the same domain. For example, the categories might be as follows: 1. Born in Utah; 2. Democrat; 3. Negro; 4. Lawyer; 5. Baptist; 6. Male. These might be categories which distinguish six different individuals or they might apply to one of the six individuals. du Mas says an attempt must be made to answer these three questions concerning both the categories or the individuals involved:

1. Along what dimension should the individuals or the categories be ordered?
2. What individual or category should have the highest rank and which the lowest?
3. How does one arrive at a score for an individual or a scale value for a category?

It is in the answering of these three questions that the

subjective element comes to the foreground. For example, the purpose for which prediction or measurement is desired will to some extent determine the rank of the individuals or the weight of the categories. Let the purpose be the prediction of success for some candidate who is to be selected or elected to a public office say school board member. If the city is in Utah, the category 'Born in Utah,' might be very important. If the school board had no member from the legal profession on it, 'Lawyer' might be the high ranking category. If the school board wishes to be thought representative and had no Negro on the board, the category 'Negro' might be important. In like manner, it could be shown that the six categories mentioned might represent the ideal qualities needed for the school board member if he were to complement the present board. If the six qualities were stressed, but no candidate fulfilled all six of them, then some decision would have to be arrived at to show what ranks or weights or values were to be given to each of the six qualities. Whether the candidates were to be elected or selected would determine how the research worker would proceed in determining the weights. Also a knowledge of which qualities were associated with which candidates would be necessary.

In the above example the ideal qualities when arranged in the order of their importance would be the manifest variable or the manifest structure. du Mas has set forth eight postulates which serve as the rules under which he is working. The first

one, "There are universes such that members of a particular universe may be allocated to a set of categories and the categories making up the set are differently associated with a continuum of magnitudes," is the major postulate. In example given above, the universe would be the people of the community who might be eligible or be chosen to run for school board member. The particular universe might be the candidates who have had their names put on the ballot or placed before the mayor or the city council. The set of categories might be the six qualities--Born in Utah, Democrat, etc. The ranking of these categories from 1 to 6 would be their association with the continuum which would be different according to the situation--time, place, composition or membership of the present school board, etc.

The postulates II through VIII are minor and will be paraphrased or explained briefly. Postulate II says that members of a particular universe may belong to more than one category. One school board candidate might belong to the categories 'Lawyer,' 'Male,' and 'Baptist.' Postulate III says that members of a universe may belong to categories which are qualitative but measured such as I.Q. or income. Postulate IV says that members of a universe may belong to categories which are qualitative and unmeasured (such as Baptist) and qualitative and measured (such as Weight). Postulate V says that a category may contain several members with the same value--say two men with the same weight 180 pounds. Postulate VI says that a category may have individ-

uals who have different values such as 160, 172 or 180 pounds even though they may all be White, Democrats, and Lawyers. Postulate VI says that even in qualitative categories which are measured two individuals may have the same value 160 pounds in this one category but not belong to the same other categories. One person of 160 pounds might be Female, White, Lawyer while another 160 pound candidate might be Male, Negro, Baptist. Postulate VIII says that individuals may be unlike in the qualitative unmeasured (male or female) and unlike in the qualitative measured (180 pounds or 140 pounds) and still be placed along the same continuum.

In brief, du Mas is saying that Manifest Structure Analysis can be used on any data. He does say, however, that someone will have to assign rank or weight to either the categories or the individuals in order to have a manifest variable. In order to find anything, one must know what he is looking for. du Mas uses his figures<sup>1</sup> called the Model and the Empirical Analogue to help the person to rank or give value to his data--to get a clear picture of his ideal or his manifest variable. If the person cannot rank or give weight to either his individuals or his categories, he can get some help from du Mas' article, 'Behavioral Scaling of Personality Tests' (5). However, this matter will not be discussed in this research paper.

<sup>1</sup>See Appendix VIII.



The word "Manifest" has two aspects, one involving subjective clarity and the other objective clarity. The first aspect was discussed in connection with the research person's having a clear idea of what he was searching for. This was a clear mental picture of the individual with the necessary qualities or qualifications as listed in the example of the school board candidate mentioned previously. The other aspect--the objective clarity is involved in visual representation by patterns, diagram, graphs, matrices, etc. The use of such visual material is explained by E.A. Fisher (7) thus: "The preliminary examination of most data is facilitated by the use of diagrams." Large back drops upon which the data can be mounted have the effect of magnifying each detail when the observer is close or of giving a general picture when the observer moves away. du Mas uses what he calls a "scaling frame," a mechanical device, in order to let the observer see the details plainly yet get a view of the whole picture.

The scaling frame can be of any size. In his book (4) du Mas describes a frame 55 inches square which would have space for one hundred individuals and one hundred categories or ten thousand cells with holes for the thumb tacks. The one which the writer saw in June 1957 at the University of Montana had 3 sections with ten thousand cells each and measured about five feet by fifteen feet. These frames are made of wood and have one hundred (or more) vertical slats about 3 1/2 inches wide each with one hundred cells in it. The frame itself looks much like a curtain

stretcher except for the grooves or troughs for the slats to stand or move along. By now an even larger frame will probably be in use for a summer workshop of 1958.<sup>2</sup>

To mount the data on the frame, the person doing the experiment must give all his individuals numbers and all the information must be classed into categories which also have numbers. A strip of paper with the individuals' numbers is tacked up along the left hand side vertically from top to bottom to correspond to the cells. Each slat is given its category number. Then the thumb tacks are

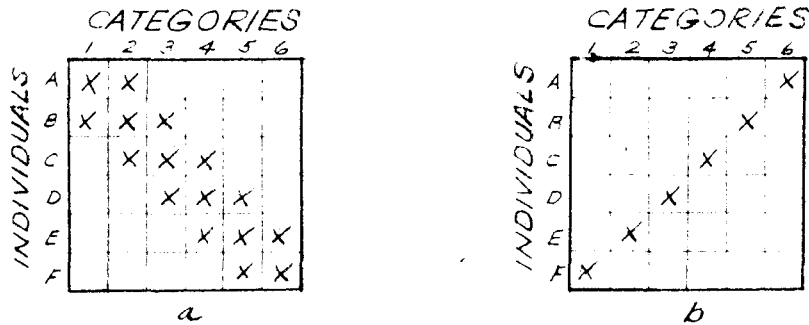


FIGURE 3. EXAMPLES OF SEGMENTAL CATESCALES.

put in to represent to which categories each individual belongs. The result at first seems like an unassembled jig saw puzzle. Then the experimenter determines empirically which of du Mas' three patterns will fit the data.

The most familiar pattern, at least to the makers of scattergrams, is the segmental catescale (the segmental pattern). This is a wide (Figure 3a) or narrow (Figure 3b) bank of thumb tacks extending diagonally from the lower left to the upper right hand

<sup>2</sup>Appendix I.

corners of the frame. It could also be a band from the upper right to the lower left if the data were all of the negative kind<sup>3</sup> or if the individuals were ranked from the lowest to the highest. By such a pattern the experimenter is really selecting the data which characterizes each individual, making him unique or different from the others. These unique features receive the weight or the rank of the person or persons having them. The category is given a value equal to the mean of all the weights or scores of the persons which belong to it.

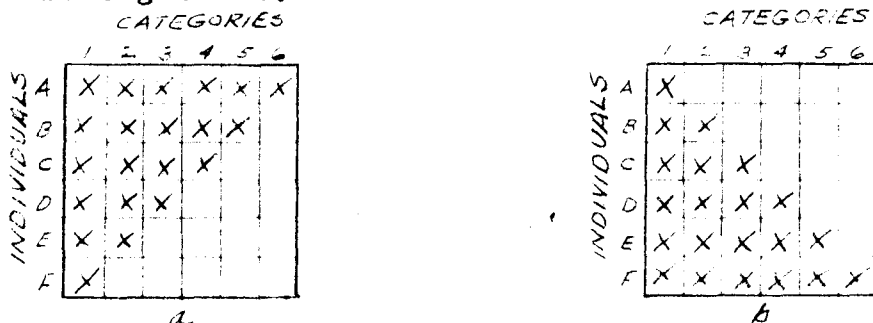


FIGURE 4. EXAMPLES OF INTENSIVE CATESCALES.

The second pattern called the intensive catescale is also a familiar one. It is a right triangle in the upper left (or lower left) corner of the scaling frame with the diagonal from the lower to the upper corners forming the hypotenuse (Figures 4a and 4 b). The thumb tacks fill in the triangle more or less solidly. If the triangle (Figure 4a) made by the filled rows of thumb tacks is in the upper left corner, the filled rows usually

<sup>3</sup>In clinical data, the negative kind might mean those children with physical defects, or social disabilities, or mental inadequacies.

represent the many assets or gifts or talents possessed by the highest ranking individuals at the top of the frame. The number of these assets, talents, etc. dwindle to a very small number in the case of those at the bottom of the scale. If the triangle (Figure 4b) made up of the filled rows of thumb tacks appears at the lower left, then it is usually found that the numerous thumb tacks at the lower part of the frame are disabilities or lack of talents or liabilities suffered by the lowest group.

		CATEGORIES									
		1	2	3	4	5	6	7	8	9	10
INDIVIDUALS	A	X	X	X	X	X		X			
	B	X	X	X	X	X		X			
	C			X							
	D	X	X		X	X					X
	E		X			X					X
	F	X		X		X				X	

FIGURE 5. EXAMPLE OF A CLUSTERY CATESCALE.

The third pattern Figure 5 called the clusterly catescale, is not a familiar one. As the word 'clusterly' implies that there will be a cluster or a grouping, this pattern cannot be made unless the individuals form groups which are internally alike in criterion scores and in categories and unlike other groups. Each person gets the score which is found for the whole group. Thus, for prediction this method seems to give less exact or discriminative information. The clusters will merely help to predict to which group a person will belong.

In order to evaluate or test any theory or method it is necessary to know what the author or creator of the theory is trying to achieve or accomplish, how he is trying to achieve his

purpose, how his procedures are different or original or more effective than any others. In his book, du Mas has set forth some restrained claims for his theory and method. However, he expresses these ideas more effectively in the material which advertises his 1958 summer workshop.<sup>4</sup> Ten of these claims as to what manifest structure analysis will help the research person to do are as follows: 1. to construct ability, aptitude, achievement, interest, or personality tests; 2. to reduce the number of items in any already prepared tests; 3. to reduce the time spent in giving, marking, or analyzing the test; 4. to evaluate test batteries, profiles, or psychographs; 5. to predict a qualitative or quantitative criterion; 6. to utilize case histories or application forms as measuring instruments; 7. to save time, space, equipment, personnel, and money in analyzing data; 8. to do more research than is now done on the present budget; 9. to maintain the present research output on a reduced budget; 10. to use manifest structure analysis on a day to day evaluation of people.

The present research has not tried to deal with all these ten claims. It has concentrated on claims 2, 3, 5, and 6. It also has given some attention to the ease of using and the simplicity of manifest structure analysis. du Mas claims that no special background in mathematics or statistics is needed in order for a person to learn his system or method. There was a

<sup>4</sup>Appendix I.

time when, in the writer's opinion, du Mas thought his book would be an adequate teacher of his method (and even his theory). The 1958 workshop is, in a way, an admission that the book was not a self-sufficient way to introduce the theory and method of manifest structure analysis.

Except for the present research paper, a Master's Thesis (11), and a review of the book in Educational and Psychological Measurement, Volume XVII Part 4 (Winter 1957) pages 634 to 636 by David G. Hays, there has been little done or written in evaluation of du Mas' contribution. At the time of this writing another evaluation to be written by S.T. Mayo is known to be in preparation for publication in Educational and Psychological Measurement in the near future. On the whole, comment from du Mas' contemporaries is conspicuous by its absence.

The strong claims and the recency of the publication of the theory and method, Manifest Structure Analysis, should lead to the critical evaluation by contemporaries. There is enough logic in du Mas' theory that it cannot be ignored or explained away, it must be evaluated. To eliminate bias the evaluation should be done by an outsider (not a student in du Mas' classes). The present paper is a one "outsider's" attempt to evaluate some of the claims for the theory.

can be trained to use the method, is to be evaluated, it can be said that a person could be trained to use the method without understanding the theory. The writer tried out various phases of the work on the following variety of people: (1) a ten year old fairly intelligent elementary school pupil; (2) an exceptionally bright high school sophomore; (3) a high school freshman of average intelligence but exceptional artistic ability; (4) a housewife with merely an elementary school education but exceptional artistic ability; (5) a housewife with little formal education but versatility in the use of business machines; (6) a housewife with 3 years of college education; (7) a college sophomore majoring in engineering; (8) an electronic engineer; (9) a statistician; (10) a high school mathematics teacher; (11) a psychologist heading an important project; (12) a specialist in tests and measurements; (13) a psychometrician. Of those who helped, there were some groups with several people in them and others with only one person. In arriving at the patterns, the artistic people and those with engineering training were most helpful. The people with the most education were on the whole less helpful than those without much education. In dealing with the mathematical computation and the use of the calculating machine, the college student, the high school sophomore, and the housewife with training in business machines were most helpful. The grade school pupil, both high school students, the college student, and all the housewives were all adept at mounting

a variety of information was collected concerning each of the children. Several experts were involved in compiling this record folder as well as many child study psychologists, adjustment teachers, elementary class room teachers, and the members of the Cooperative Research Project staff.

Because the writer felt that the du Mas scaling frame was one of the weak points in his system, she devised a new scaling frame. The writer felt that the du Mas scaling frame (even the small 55 inch square one made of plywood) was too clumsy and too heavy to carry around or to store easily or to hang up on the wall conveniently. The large 15 by 5 foot frame was not only very difficult to move around but practically had to have a special room for its storage and use. Beside the above disadvantages, the scaling frame was expensive to have made by professional carpenters<sup>4</sup> or required space, tools, skill, and time if the person were to make it himself.

The scaling frame which the writer adapted to a new usage was from a commercial product called the Acme Visible File (largest size). The file when closed looks like a large book with covers made of light weight aluminum. Inside the covers along the edges from the top to the bottom are grooves. The pockets which are attached to hinge strips are moved along these grooves. The pockets are equivalent to the slats in the du Mas scaling frame.

<sup>4</sup>Appendix VI.



They are about as easy to remove from the groove as are the du Mas slats. Each pocket will take care of 52 individuals. du Mas' slats take care of 50 or 100. Ordinarily each cover holds 49 pockets but 65 can be fitted into the groove. By using both covers 130 categories and 104 people or 13,520 cells are available. On du Mas' 55 inch scaling frame, there are 100 people and 100 categories or 10,000 cells available.

The advantages of this scaling frame are: 1. its size--21 by 14 inch when folded closed or 21 by 28 inch when open; 2. its weight--6 1/4 pounds with its standard 98 pockets; 3. its cost--\$42.05 (if no extra pockets are purchased) to \$138.80;<sup>5</sup> 4. the ease and speed of acquiring (can be ordered and delivered in a week); 5. the possibility of keeping a record of every category by merely taking the title inserts from their casings. This last advantage is a worthwhile one if two people were desirous of using the scaling frame at the same time. One would merely remove his title inserts and let the next person slip his in. This could never be done on the du Mas frame where thumb tacks are used and will fall out if the slats are not moved carefully or stacked along side of (not on top of) each other.

The visible file has some disadvantages. It is not large enough to show clearly the details as does the large du Mas scaling frame at the University of Montana. The cellophane strips

<sup>5</sup>Appendix V.

under which the title inserts and the markers (or flags) are slipped have a sheen which makes them sparkle or glare with reflected light. This is disconcerting at some times and to some people. The envelopes are fastened to hinge strips which are 13 1/2 inches long, 1/4 inch wide, and 1/8 inch thick. These sometimes break off and have to be replaced as the file is used frequently. Replacements are easy to procure and the pockets stapled on to them. These strips sometimes do not move easily in the groove and have to be sandpapered. If both covers of the file are to be used, they have to be unhinged and placed so that the 21 inch sides are next to each other without the 3 inch piece which contains the hinges. Even without this 3 inch piece there is a line where the two sections come together.<sup>6</sup> Markers or flags sometimes slip out of place and after experimenting with using marking ink or using flags, the writer decided marking ink was cheaper and more satisfactory.

It is not easy to compare the two scaling frames but the price range for the du Mas frames would be from \$40.40<sup>7</sup> to \$2,646.20.<sup>8</sup> The range for the Acme Visible file would be from \$42.05 to \$133.30.<sup>9</sup> The weight range for the du Mas scaling frame would probably be from 10 to 100 pounds. The Acme file

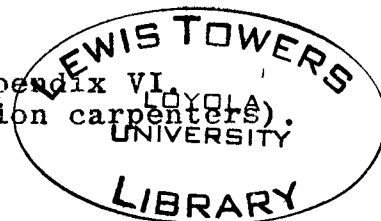
<sup>6</sup>Appendix IX, X, XI, XII, XIII.

<sup>7</sup>Cost of material; no labor costs.

<sup>8</sup>Cost of material plus labor costs (union carpenters).

<sup>9</sup>Appendix VI.

Appendix V.



weight range is between 6 1/4 to 10 pounds.

There was only one other special device which the writer used. In connection with the clustery catescale the du Mas method of the filled row technique or the index of communality were used to tell to which group a person belonged. Both of these methods were excellent if there were a small number of groups and very few categories in each group. The writer had a small number of groups (13) but a large number of categories (41). How these 13 groups each with its filled or unfilled categories were mounted on a cardboard is shown elsewhere in the dissertation.<sup>10</sup> Each individual who was to be placed in a group had a strip of paper with the same 41 categories. His strip of paper had the categories to which he belonged marked off with the same marking ink as had been used on the 13 groups. By comparing each person with each group his points of agreement both for filled and unfilled cells could be counted and the ratio of communality could be given in fraction form. The person belonged to that group for which he had the highest ratio of communality (largest fraction). A detailed discussion of the index of communality is found in Chapter Four.

<sup>10</sup> Appendix VII.

## CHAPTER IV

### APPLICATION OF THEORY AND RESULTS

Having been given the permission to use the Cooperative Research Project's case record folders but not being permitted to remove these records from the location of the Project, the writer's first task was to find out the range and the distribution of the reading scores. Since du Mas had stressed, in a personal talk with the writer, the importance of getting a rectangular sample for the use in building the catescales, the writer tried to get two or more cases for each reading score. It was not possible even to get one case for each score from 1.0 to 6.1. As explained in Chapter II, there were no cases with scores 5.3, 5.4, 5.6, 5.8, 5.9, and 6.0 and there was only one case each for scores 1.1, 4.3, 5.0, 5.2, 5.7, and 6.1. Because the staff of the Cooperative Research Project had already decided against collecting the same data for new pupils tested as they had for their original 270 cases, there was no possibility of getting any cases with the missing scores (5.3, 5.4, etc.) mentioned above.

The first decision, that of getting a rectangular sample, had to be made, and there was nothing in du Mas' book (4) to help with the practical decision. The problem was to get a rectangular sample from data which were not adequate. There seemed to be four possible solutions. First, the data could be grouped. Second, there could be one sample for each score from 1.0 to 5.2 and scores 5.3 to 6.1 could be left out. Third, these could be

two samples for each score from 1.0 to 4.2 and the scores above 4.2 could be omitted. Fourth, every case could be left in and two cases for each score could be used as frequently as the data allowed.

TABLE I

DISTRIBUTION OF 264 CASES					
SCORE	NO. OF CASES	SCORE	NO. OF CASES	SCORE	NO. OF CASES
1) 1.0	58	19) 2.8	8	37) 4.6	0
2) 1.1	1	20) 2.9	9	38) 4.7	2
3) 1.2	2	21) 3.0	9	39) 4.8	2
4) 1.3	10	22) 3.1	7	40) 4.9	2
5) 1.4	5	23) 3.2	7	41) 5.0	1
6) 1.5	5	24) 3.3	7	42) 5.1	2
7) 1.6	6	25) 3.4	10	43) 5.2	1
8) 1.7	7	26) 3.5	3	44) 5.3	0
9) 1.8	9	27) 3.6	8	45) 5.4	0
10) 1.9	6	28) 3.7	2	46) 5.5	1
11) 2.0	7	29) 3.8	4	47) 5.6	0
12) 2.1	4	30) 3.9	5	48) 5.7	1
13) 2.2	5	31) 4.0	5	49) 5.8	0
14) 2.3	3	32) 4.1	4	50) 5.9	0
15) 2.4	5	33) 4.2	9	51) 6.0	0
16) 2.5	6	34) 4.3	1	52) 6.1	1
17) 2.6	4	35) 4.4	2		
18) 2.7	7	36) 4.5	0		

In order to get a rectangular distribution, two conditions must be satisfied: (a) the intervals must have the same number of equal units and (b) the same number of cases. The units in this study might be called tenths of a grade. One individual with a reading score of 6.1 would be a somewhat more efficient reader than a child with a score of 6.0. The difference between 6.1 and 6.0 is one-tenth of the 6th grade expected progress in reading efficiency. The range of scores from 1.0 to 6.1 contains 52 divisions representing tenths of a grade. Although there were 270

cases listed as studied, there were only 264 which were adequate. Six were incomplete or inadequate for some reason or other.

These 264 cases were distributed as shown on the previous page.

In these 52 cases there could be 13 divisions with 4 scores in each such as 1.0-1.3, 1.4-1.7, etc. or 4 divisions with 13 scores in each such as 1.0-2.2, 2.3-3.5, etc. or there could be 52 divisions containing 1 score each. The object was to get as large a rectangular sample as possible. The determining factors were the upper set of scores, that is the scores 4.2 and larger.

TABLE II

DISTRIBUTION OF 264 CASES IN THIRTEEN GROUPS

SCORES	NO. OF CASES	SCORES	NO. OF CASES	SCORES	NO. OF CASES
1) 1.0-1.3	71	6) 3.0-3.3	30	11) 5.0-5.3	4
2) 1.4-1.7	23	7) 3.4-3.7	23	12) 5.4-5.7	2
3) 1.8-2.1	16	8) 3.8-4.1	18	13) 5.8-6.1	1
4) 2.2-2.5	19	9) 4.2-4.5	13		
5) 2.6-2.9	28	10) 4.6-4.9	6		

With 13 divisions the arrangements would be as shown above. If the number of the interval was thus determined, the number of cases in each group would supposedly be set by the number of people in interval 13, the smallest group. There would then be 1 person from each interval or 13 people would be the size of the sample. If the 12th interval were selected, there would then be 12 intervals with 2 persons and one interval with 1 individual giving a total sample of 25 individuals. If the 11th interval were selected, there would be 11 intervals with 4 individuals, one interval with 2, and one interval with 1 individual--a total of 47 persons in the sample. If the 10th interval were selected,

there would be 10 intervals with 6 individuals each, one interval with 4, another with 2, and one with 1, total being 67 persons.

This was coming near to the desirable size for a sample, namely 100 individuals, but the rectangular distribution was not being maintained in 3/13th or 23 per cent of the intervals.

Now if four intervals were set up, the arrangement would be as follows:

TABLE III

DISTRIBUTION OF 264 CASES IN FOUR INTERVALS

SCORES		NO. OF CASES	SCORES		NO. OF CASES
1)	1.0-2.2	125	3)	3.6-4.8	45
2)	2.3-3.5	85	4)	4.9-6.1	9

If the smallest interval (4) with 9 cases were taken, then 36 cases would comprise the sample. If the next largest interval (3) were taken, then there would be 3 intervals of 45 individuals each and 1 interval of 9 totaling 144 persons in the sample. This was an adequate sample, but it left only 124 for validating the results and would not maintain a rectangular distribution in 25 per cent of the intervals.

If the number of intervals were left at 52 and the number of cases were set at 2 for each interval, there would be 37 intervals with 2 each and 7 intervals with 1 each giving a total of 81 cases. There was also 1 case 4.4 which could just as well be classified as 4.3 and an extra 3.6 which was put in by mistake. Although this distribution did not give a sample of 100 cases and although it violated the ideal rectangular distribution in 15/52

or 29 per cent of the intervals, nevertheless it was chosen as the best arrangement for getting finer discrimination between scores. The group left for validating purposes was large in number but not in range. This was excellent for our purposes except for the large number of 1.0 scores.

TABLE IV  
GRADE PLACEMENT

1.	High School	(Regular)	1	} 31
2.	"	(Ungraded)	0	
3.	Adv.	Elementary	30	
4.	Lower	"	17	} 12
5.	8A Regular	"	3	
6.	8B	"	3	
7.	7A	"	4	
8.	7B	"	2	} 10
9.	6A	"	3	
10.	6B	"	2	
11.	5A	"	3	} 9
12.	5B	"	2	
13.	4A	"	2	
14.	4B	"	5	} 4
15.	3A	"	1	
16.	3B	"	1	
17.	2A	"	2	} 4
18.	2B	"	0	
19.	1A	"	2	
20.	1B	"	0	} 0
21.	1C	"	0	
22.	Kindergarten		0	

Since the first problem had been settled, the data from these 83 case folders was carefully copied from the record folder.<sup>1</sup> There was a grouping of sorts on the record folder but it seemed awkward in places.<sup>2</sup> For example, in the table above if all the B

<sup>1</sup>Appendix IV.

<sup>2</sup>12-1 B Semester, 12-2 A Semester, Appendix IV.



semesters were to be classed together there would be 15 cases ranging from 1B to 8B. The A semesters would have 18 cases ranging from 1A to 3A. There were 22 classes for this data as shown in the table. A reduction in the number of groups was made by placing in one group the advanced ungraded and both high school groups, the lower ungraded in the second, the 8A to 7B in a third, 6A to 5B in a fourth, 4A to 3B in a fifth, 2A and below in a sixth. As was shown in this one case, grade placement, 22 groups or categories were reduced from 22 to 6 and later to 4. For each question or item in the Case Record Folder (Appendix IV) there were 3 to 50 answers or descriptive phrases. These nearly 2000 answers or descriptive phrases were reduced to 275 categories and later to 152 categories.

When the 152 groups were determined, then they were numbered and called categories. Each of the 83 cases had all data classed into the proper categories and the category numbers listed. Each pocket of the visible file was equipped with a title insert containing a category number at the top and 83 numbers following below it to represent the 83 individuals. Each of the 83 individuals was ranked according to his reading score. Number one had the highest reading score, 6.1 and 82 and 83 had the lowest score, 1.0. Little green signals or green marking ink were used to tell which of the 83 persons belonged to the category which was being prepared at the time. When all 152 categories were thus marked, the pockets were put into the Visible Filing Case

or Frame<sup>3</sup> and a picture taken. The picture has been enlarged to show the lack of pattern and the general disordered, unorganized appearance of the data.<sup>4</sup>

At this point the instructions (4, p. 68) said to reject non-scalable material on the following bases: 1. Multimodal; 2. Gappy; 3. Associated with a large part of the range; 4. Not sufficiently associated with individuals in the sample. These four bases for rejecting non-scalable material seem adequate on the surface but are not really helpful. Multimodal would seem to mean a category which had more than two modes. Guilford (9, p. 63) defines a mode as "the point on the scale of measurement with maximum frequency in a distribution." Now since our visual presentation would not show any piling up of scores at a point unless the whole category is thought of as a point, this definition is not useful. A mode in a category would be a long line of cells filled with thumb tacks or colored signals. The longest solid (filled) line would be the mode. If the line of signals designated as the mode had 17 individuals whose ranks were consecutive from the 3rd to the 19th, could another group of 16 individuals whose ranks were consecutive from 68 to 83 be a lesser mode? According to Guilford's definition, it could not; according to practice and du Mas' purpose, it could be. This

<sup>3</sup>Only 65 could actually be put in the channel; the rest had to be placed along side.

<sup>4</sup>Appendix IX.

category would then be bi-modal. Should it be rejected as multi-modal? du Mas' book does not answer such a problem. In the segmental catescale a band with a width of 17 individuals is really too wide, and the writer would reject it.

Having found no real help on the first basis, the writer next looked at the second basis for rejecting catescales, the gappy criterion. A general definition for 'gappy' would be a continuum with sections, divisions, or steps which has had openings left by missing sections. How many openings in a continuum or a catescale would be needed for it to be called gappy? Hasse (11) uses a rule of thumb "3 to 1." This seemed to mean that 27 or more openings (not consecutive) would make a gappy distribution in a group of 83 individuals. The writer felt that many less than 27 would make a distribution gappy. The number of missing sections, the writer felt should be two or more not one or more.

The third basis, associated with a large part of the range, was logical, but no indication was given as to how many individuals could be associated with a category before the category became useless for the segmental catescale. The writer decided this question by thinking of the wide band pattern. If there were eight to ten individuals consecutively all belonging to a category, this would make a very wide band and any other belonging to the category would have to be few in number and have gaps of eight or ten between them and the group making up the band.

The fourth basis seemed most confusing for, if the grouping

was well planned, there would have to be two or more individuals in each category. If these two or three were consecutive--such as 81, 82, 83--the category would be very useful for the segmental cateScale. However, if these three individuals were scattered in such a way as to be 1, 36, and 80, then the category would be rejected.

When the process of rejecting non-scalable material had reduced the number of categories to about half its previous size, the writer started to rearrange the material so as to form the diagonal band (the segmental scale pattern). The writer sorted out the categories so that all the categories which had a concentration of markers (flags or signals) at the top were placed together; those which had a concentration in the middle were put in another group and those which had a concentration at the bottom in the last group. This rearranging could be done in another way. As was mentioned before each category had been given a weight or value by adding up the scores of the persons belonging to it and dividing by the number of individuals. The categories can be arranged in ascending or descending order by putting these category weights in either ascending or descending numerical order. If these weights have not been distorted by a few extreme cases, the average will show mathematically about where the concentration of markers will be. In trying to get a clear cut pattern, there will be a further rejecting of material.

The writer had a feeling that because her pattern<sup>5</sup> was not really a clear cut band, she had not rejected enough non-scalable categories. However, in order to get a fine looking diagonal, the number of categories and the number of individuals should be equal. du Mas does not say this but diagrams of his and others usually depict a square with the individuals and categories equal in number. Also du Mas' original scaling frame was a square. The segmental catescale was finally set up with 26 categories. There were probably too many categories rejected because in the application of weights to the 181 cases, there were 3 cases which could not be given a value because they belonged to none of the 26 categories.

As was mentioned before a score called an S (derived) score was calculated for each individual by the formula  $S = \frac{1}{N} (\sum V)$ . In the formula N was the number of categories to which the individual for which the score was calculated belonged.  $\sum V$  was the summation of the values of the categories. The means of the R (reading scores) distribution and the S distributions were 3.1 and 3.03. Although the standard deviations and the means were calculated for each distribution the writer could not "tell by inspection whether or not a transformation of S is required and also whether or not the correction, K (the difference between the means of the two distributions) should be made for the S

<sup>5</sup>Appendix X.

values." (4, P. 71) Not having the necessary background of experience to make such judgments, the writer used the K (.07) and figured out a new set of values S'. Because these did not bring the R and S' distributions closer together or more alike and because it seemed to distort the situation with 181 cases, the writer discarded the plan of adding +.07 to all the S scores in order to get a new S' distribution. The null hypothesis was set up that there was no significant difference between the two means. The standard deviation for the R distribution was 3.09 and for S was 3.04. The standard error of the mean of R was .3392 and for S was .3337. The standard error of the difference was calculated to be .26. With an actual difference of .07, a critical ratio of .27 was obtained. Therefore, the null hypothesis was not rejected, and there was no need for a transformation of S. This was done according to the method suggested in Garrett (8, pp. 213-216).

The next step in the instructions for extracting a segmental catescale was to calculate the product-moment correlation  $r$  between the paired values of R and S. du Mas (4, p. 71) had a correlation in his example of +.99 but the writer was able to achieve a correlation of only +.70.

According to step eleven (4, p. 74) a Chi square test of the two distributions should be made. The theoretical or expected frequency would be the frequency of the R values and the frequency of the S values would be the observed frequency. Diffi-

culties arose here because the S distribution had scores with two decimal places while the R distribution had scores with one decimal place. When the S distribution scores were rounded off, then there were many S scores which were alike. The R scores were of rectangular distribution with only two scores alike in most of the cases. In the example in the book (4, p. 74) the Chi square test showed the distributions to be similar. The writer calculated a Chi square of 227.33 with 39 degrees of freedom. Since the tables did not go beyond 30 degrees of freedom, the writer used the formula  $t = \sqrt{\chi^2 / df}$  (9, p. 540) converting the Chi squares to a t ratio of 13.6. From the t table (9, p. 539) and (8, p. 427), it seemed reasonable to reject the null hypothesis. It was concluded that the distributions were not similar. du Mas in his Chi square test mentioned above had shown his distributions to be similar.

Since it was necessary for the distributions to be similar, the writer went back to the distribution of scores made by adding .07 to each of the S scores to see if this change would help with the Chi square test. The Chi square for this distribution was 224.16 with 41 degrees of freedom. By conversion to the t as above, the t was found to be 12.17 which, as was found above, was well beyond the one per cent level of significance. The null hypothesis had to be rejected and the distributions still could not be considered alike.

The writer's next effort to get an S distribution similar to

R was to use the formula  $S' = MS + K$   $M = \frac{N \sum RS - (\sum R)(\sum S)}{N \sum S^2 - (\sum S)^2}$  and  $K = \frac{\sum R \sum S - \sum S^2 \sum R}{(\sum S)^2 - N \sum S^2}$ . By the use of this formula a new set of scores  $S'$  were calculated. The Chi square test for these gave a Chi square of 67.66 with 36 degrees of freedom. By converting to  $t$ , the  $t$  was found to be 3.2. This was beyond the 1% level of significance, so that the null hypothesis was rejected.

As it was suggested that the Yates correction for use when table entries are small would be suitable for this data, the next step was to go back over the Chi square calculations and make the Yates correction. For the first Chi square test, the one in which Chi square was equal to 227.33, the use of the Yates correction changed the Chi square to 177.666 or 177.67. However, when this was changed to a " $t$ ," the " $t$ " was equal to 10.075 which was beyond the .001 level of significance. The null hypothesis would have to be rejected and the distributions would be considered as unlike. The second set of  $S$  values (ones which .07 had been added), the Chi squares changed from 224.16 to 167.375 with the use of the Yates correction. When a " $t$ " of 6.32 was found and looked up in a table (15, p. 248), the number was found to be beyond the .001 level of significance. The null hypothesis was rejected and the conclusion drawn that the distributions were still unlike. When the Yates correction was used on the  $S'$  distribution which had given a Chi square of 67.66 with 36 degrees of freedom, the Chi square was 39.083. By extending the Chi square table to 36 degrees of freedom from the 30 at



which it stopped, the writer calculated that the level of significance was .20. Therefore, the null hypothesis was not rejected. The distributions might possibly be considered similar. The conversion to a "t" score gave a value of .415. An answer as small as this with 36 degrees of freedom implied a probability larger than .10. The probability might well be the .30 calculated above. However, since anything above or larger than .05 could be considered suitable evidence for not rejecting the null hypothesis, the null hypothesis was not rejected and the distributions were considered to be alike. A product-moment correlation was calculated for this new S' distribution and was found to be .703.

Next the values were applied to the 173 cases.<sup>6</sup> A product-moment correlation was calculated and found to be .524. There were some cases with data missing. If these cases are omitted, a Chi square test for the 173 cases gave 605.408 for 32 degrees of freedom with the Yates correction used. The t was found to equal 26.859--well beyond the one per cent level of significance and the null hypothesis had to be rejected.

As each catescale was built, photographs were taken and enlargements were made. These enlargements are found in the appendices X, XI, XII, and XIII.

<sup>6</sup>There had been 181 but 3 were not able to be evaluated for they had none of the characteristics chosen.

These statistics seem to show that although the writer did not get the very high correlation of .98 or .99 which du Mas had (4, p. 71) between the R (criterion) score and the S or S' (derived) scores nevertheless .703 can be considered as denoting a high relationship (8, p. 173). When the ordinary weights were applied to the 178 or the 159 cases the correlation was such as to be classified as substantial or marked. Contrary to expectation, when  $S' = MX + K$  was used on this distribution, it lowered the

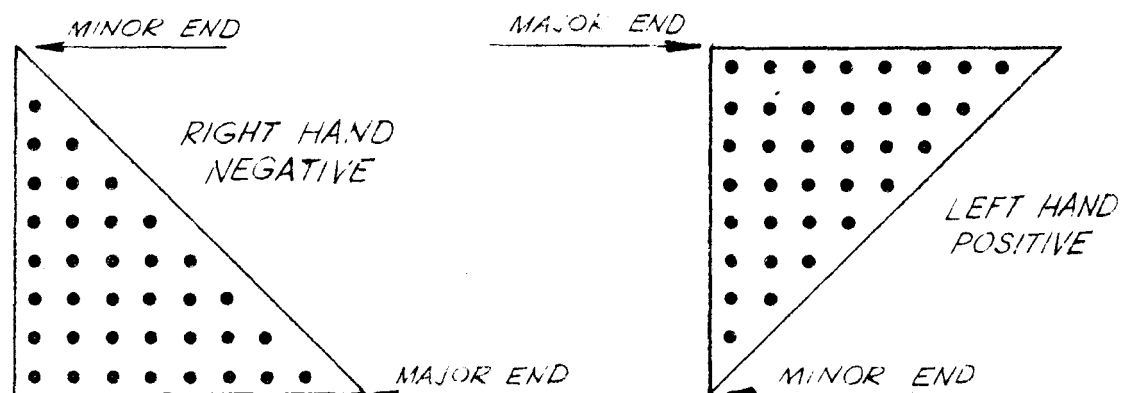


FIGURE 6. EXAMPLES OF MAJOR AND MINOR ENDS OF ASSOCIATION SURFACES.

correlation from .524 to .476. This would seem to imply that the sample used to build the catescales was not representative.

The next step was to start building the intensive catescale with its two aspects. The procedure was very similar to that done for the segmental catescale until the rejection of non-scalable material was started. According to the directions (4, p.91), some of the bases were the same namely, multimodal and gappy. There were two new bases--not anchored at the major end of the association surface and low associated categories. Since there

were no explanations of these, only examples, the writer will give her own interpretation. The major end of the association surface would mean to the writer, the end which will have the large number of cases. The minor end will mean the vertex or the end with the very few cases. It does not take too great powers of visualization to see that if the category had no persons at the major end (either top or bottom depending on which variation was attempted), the pattern would be reduced to a wide band or segmental catescale. The fourth basis seemed to be a mere variation of the wording of the fourth basis of the segmental catescale. This meant to the writer too few individuals belonging to the category and the position of these few was not close enough together.

Since this pattern with both variations (see Figure 6 on previous page) seemed the easiest and most effective to achieve, the writer has assumed that the material best adapts itself to this pattern. This does not always occur. Hasse (11) could not achieve either variation of this catescale and hence omitted it entirely from his report. The right pattern was built easily for these cases of mentally handicapped children who were often physically, educationally, emotionally, financially, and socially handicapped as well. There were 53 categories selected to make this pattern, and so this perhaps shows the writer's judgment had improved with experience.

The score was arrived at by the same formula<sup>6</sup> but this score was not sufficient as was the case in the segmental scale. R will not be expected to equal S but S'. S' was to be found by the formula  $R=MS+K$ . The formula is the slope-intercept formula from analytical geometry with M the slope and K the R intercept. The formula for finding M (1, p. 93) was  $M = \frac{N \sum RS - \sum R \sum S}{N \sum S^2 - (\sum S)^2}$  and for K,  $K = \frac{\sum RS - \sum R \sum S^2}{(\sum S^2) - N \sum S^2}$ . The M is equivalent to the usual regression coefficient for one criterion variable and one prediction variable. After computing the S' values as suggested by the above formulae, the product-moment correlations were found between R and S' to be .773.

When the test weights were used on the 181 cases, there were no cases which had to be omitted because they lacked all of the 53 categories. The product-moment  $r$  was .559 when all 181 cases were used. There were cases with data missing which might have contributed to a higher correlation between the R and S' values. When these cases were omitted, there were 126 cases and the correlation rose to .575, or .58.

Although the writer could have stopped with this pattern as a test of the intensive catescale, she photographed the negative variation<sup>7</sup> and proceeded to build the left handed triangle which she called positive. In this triangle, the large numbers of categories were at the top with the high reading score individ-

<sup>6</sup> $S = \left(\frac{1}{N}\right) (\sum V)$ .

<sup>7</sup>Appendix XI.

uals. These represented assets or advantages which seemed to help or accompany high reading scores. The procedures were just the same as before but they gave an  $r_{KS}$  value of .696. There were 60 categories selected to make this pattern and when the weights were applied to the 181 cases there were no cases which had to be left out because they belonged to none of the 60 categories. The correlation was amazingly high, being .81. When the incomplete cases were left out, there were 126 cases, and the correlation was .47.

After photographing the positive variations of the intensive catescale, the writer started on the clustery catescale. All procedures were the same as before but the bases for rejection were only two in number--gappy and low association (too few and too far apart). The rearrangement of the pockets followed the rejection of non-scalable material. Of the categories 41 were retained. A tremendous amount of calculation went into the selection of these 41 categories because no patterns were visible--no groups appeared. Tentative cutting points and tentative groups were formed. For each category that was allowed to remain, the writer had to prove that it had great value to some particular group and did not add to the number of ambiguous clusters. Even after careful selection and testing, the patterns were not plain. Mathematically, these groups of individuals could be proved to be more like each other than they were like any other group by the index of communality. To explain

this as simply as possible, an example of A and B might be given. A answered questions 1, 3, 5, 7, 9, and 10 with a yes and questions 2, 4, 6, and 8 with a no. B answered 1, 3, 5, 7, and 9 with a yes and 2, 4, 6, 8, 9, and 10 with a no. The index of communality is 90 for they agree in their yeses and noes 9 times out of 10 and disagree only once. These two have a high degree of communality and belong together in a cluster. Person C answers questions 1, 2, 3, 4, 5, and 6 with a yes and 7, 8, 9, and 10 with a no. When A and C are compared their index of communality is only 40 for they agree only on yes answers 1, 3, and 5 and on no answer 8. C has a communality index of 50 when compared with B but this is not really high. However, if C were compared with D who answered yes to questions 2, 4, 6, 8, and 10 and no to questions 1, 3, 5, 7, and 9, C would also have a communality index of 50 with D. C would certainly be classed as an ambiguous cluster--no way of telling whether he was more like B or D. To decide, the writer would have to find out the communality index with E and perhaps even F before deciding whether C belonged with A and B group or the D, E, and F group. A short way might be to remove one or two categories. For example, to remove category 9 will raise the communality between A and B to 100; between A and C to 45; B and C to 55 and lower C and D to 45. Now C could very well be placed with A and B unless E and F showed very high communality with C.

It can truthfully be said that the selection of the 13 clusters was not done visually but mathematically much in the way which has been shown above--namely by calculating indices of communality and by removing categories which made for ambiguous clusters. The value of each cluster was computed by adding the R scores of all the persons belonging to the cluster and dividing by the number of people belonging to it. Each person received the S value of the cluster to which he belonged. When the product-moment  $r$  was calculated for the R and S scores it was .976.

The application of the weights to the 181 cases was very difficult, for indices of communality had to be figured out for each of the 181 cases as compared to each of the 13 clusters. The individual was given the cluster value of that cluster for which he had the highest index of communality. When the R scores were compared with the S scores, the correlation was .223 for the 181 cases and .224 when incomplete cases were left out.

In summarizing the results, the writer was able to find the best results from the two intensive patterns. A reference to the photographs will show that these two triangular patterns are clear. The other two are visually poor and so far as the results are concerned--inferior to the intensive catescales. For the purpose of prediction the clustery catescale was practically useless. It required a tremendous amount of work and had low correlation. The Chi square tests and the Wilcoxon Matched-Pairs Signed-Ranks Test showed that in testing the similarity of the R

distribution and the S or S' distributions, the two distributions were unlike. The writer wished to find the opposite and to accept the null hypothesis but could not do so except in the case of the Intensive Catescale I. All the catescales passed the reliability test of du Mas (4, p. 75) and all but the clustery catescale passed the validity test. In what the writer would call the adequacy of coverage test and du Mas the "calculating the portion of the sample for whom a score is determinate," all catescales except the segmental had a perfect score 1.00. The segmental had <sup>89</sup>~~83~~ or .96.

It is interesting to note along the line of the writer's trouble with the Chi square tests and her inability to achieve the high correlations which du Mas mentioned in his book that du Mas' latest published article (6) has cases with correlations of .72 and .82 and does not mention any Chi square tests. In this analysis, du Mas too achieved the best results from his intensive catescale and mentions that S' is used rather than S in order to increase the range. Another matter the writer noted in this work was that du Mas mentioned carrying out the computation to 5 figures and rounding off the numbers to 4 figures. The writer mentioned how much she felt was lost in the line of discrimination of scores because her S' scores had to be held down to two figures so as to agree with the R scores which had only two figure accuracy.



## CHAPTER V

### CRITICAL EVALUATION

If a general statement were to be made in criticism of the book, Manifest Structure Analysis, it would be that the book is too complex for some readers and too simple for others. For example, du Mas has not asked himself how much psychology, logic, mathematics, statistics and experience the reader must have in order to understand it. For example, on page 2 du Mas speaks of "Weber's ratio," of "parameters," of "stimulus material." It is doubtful that the ordinary school teacher really understands the ideas behind these terms. The first and third of the formulae<sup>1</sup> in the footnote look complex and are beyond the mathematical experience of the majority of teachers and administrators. The second and fourth formulae look less formidable but confuse the average teacher or college student by the use of subscripts and "prime" marks. Formulae such as those in the footnote are found throughout the book (i.e. pages 6, 40, 41, 42, 43, and 44). Even the educated person has a great feeling of insecurity as to his ability to read and understand mathematical material. A better than average background in logic, statistics and mathematics is needed to understand the formulae. Chapter 2 on "Evidence and Postulates" certainly requires a far greater background in geometry and logic than is possessed by the average

<sup>1</sup>  $\gamma(X, +1, P_{XA}'s) \geq \gamma(X, P_{XA}'s); M_R - M_S = 0; M_R - M_S \rightarrow \infty; S' = S + K.$

teacher or college student. As to examples of materials which may be too simple--page 65 shows and tells in detail about a data card. The writer felt this was too much space and emphasis on a matter which had to be decided by the experimenter when he collected and started to work with his data. Page 77 has a graphic depiction of identical measurements obtained from a single calibrated stick. This should be redundant for anyone who understands the theory behind a ruler or a meter stick. Appendix VIII is another example of material which is too easy or too hard.

If the book were to be used as a textbook for college students, then references, practice problems examples--more numerous and with greater detail would have to be provided. This is to say in another way that du Mas has not decided as to what will be the nature or calibre of the reader for whom he is writing.

Another weakness which du Mas has recognized as evidenced by his workshop (Appendix I), is that experience is needed to provide a basis for judgment. For example, in trying to achieve one of the patterns or catescales, the experimenter will have trouble trying to decide what categories should be discarded as "gappy." He will be plagued with the problem of whether to use a small amount of material and have a perfect rectangular distribution or to use a larger amount of material and to have an imperfect distribution. He will have difficulty deciding whether he will reduce the number of categories to a small number so as to get a very clear pattern and thereby run the risk of being unable to have an S score for certain persons who happen not to

belong to any of the chosen categories; or whether he will keep the larger number of categories in the pattern even though they reduce the correlation and obscure the pattern. The same weakness might be expressed in another way by saying that du Mas has not foreseen the problems which his readers might encounter and has not adequately provided for the solution of these problems.

In summarizing the main weaknesses of the book, it may be said that du Mas has not analyzed his public (readers) carefully enough, or has not written for a particular public, or has not foreseen the need for experience which will provide the basis for the judgment required in making decisions in the application of his theory and method.

The rectifying of these weaknesses has been begun in the holding of a summer workshop (Appendix I). Perhaps the second step will be the rewriting of the book in the form of a textbook which will incorporate examples and solutions of problems as studied in the workshop. Perhaps a workbook should be written to implement the use of the book "Manifest Structure Analysis," or perhaps a workbook should accompany the new textbook. A series of courses offered to the average college student in psychology and education would help du Mas to write at the level of the average college junior or school teacher. du Mas has the gift of expressing complex ideas in a simple and an attractive way and could write a textbook for any level he chose.

If the claim, for Manifest Structure Analysis, that anyone

can be trained to use the method, is to be evaluated, it can be said that a person could be trained to use the method without understanding the theory. The writer tried out various phases of the work on the following variety of people: (1) a ten year old fairly intelligent elementary school pupil; (2) an exceptionally bright high school sophomore; (3) a high school freshman of average intelligence but exceptional artistic ability; (4) a housewife with merely an elementary school education but exceptional artistic ability; (5) a housewife with little formal education but versatility in the use of business machines; (6) a housewife with 3 years of college education; (7) a college sophomore majoring in engineering; (8) an electronic engineer; (9) a statistician; (10) a high school mathematics teacher; (11) a psychologist heading an important project; (12) a specialist in tests and measurements; (13) a psychometrician. Of those who helped, there were some groups with several people in them and others with only one person. In arriving at the patterns, the artistic people and those with engineering training were most helpful. The people with the most education were on the whole less helpful than those without much education. In dealing with the mathematical computation and the use of the calculating machine, the college student, the high school sophomore, and the housewife with training in business machines were most helpful. The grade school pupil, both high school students, the college student, and all the housewives were all adept at mounting

material, classifying, copying data, and sorting it for various purposes. In the writer's opinion du Mas is correct in affirming that much of the work can be delegated to people with no particular education or training. However, there must be someone who must be able to make decisions and to exercise judgment.

Now as to the claim that the items of any test, case history, application form can be reduced, the writer would agree to this also. Of the 152 categories which resulted from an analysis of the answers in the record folders, the number was cut to a low of 26 categories in the segmental catescale, a middle of 41 in the clustery catescale, and a high of 59 in the left intensive catescale.

Although the writer did not achieve the high correlations between the known reading scores (R) and the predicted reading scores (S or S') which du Mas achieved, she did have correlations which ranged from .22 (clustery) to .58 (segmental). This was the range in testing the validity. When the correlations were used to test the reliability, the range was from a low of .69 (segmental) to a high of .98 (clustery).

As to the claim that the MSA method can be used on diverse kinds of material, the writer would agree. The method seems to be able to fit any kind of data. This claim is worded "Utilize case histories or application forms as measuring instruments" (Appendix I).

Besides the tests for reliability and validity which du Mas

suggested, the writer tried other tests as suggested in Siegel (16, pp. 75-83) and conferred with Dr. Itkin and the statistician of the Cooperative Research Project (C.R.P.), Mr. Brauer. Mr. Brauer felt that the conclusions shown in Appendix XV and XVI were almost identical to the findings of the C.R.P. This was interesting, because the patterns shown in these appendices were the left and right intensive catescales--the clearest of the patterns to be evolved. There were points of difference such as answer 6 in Appendix XVI that is, all other forms of intelligence tests than the Stanford-Binet. At first the statistician believed that all the children had been given the Stanford-Binet. The sociologist reminded him that the older pupils had been given the Wechsler Intelligence scale for children and the Wechsler-Bellevue for adults, while the hard of hearing had been tested by another test. This name or nature of the test had been a discriminating factor which their method analysis had not shown up. Item 19, that is, not placed because of unavailability of placement, and item 20, that is, not placed because placement not urgent, were not shown as important in the C.R.P. method of analysis. However, these could be discriminating items. The items with an asterisk (\*) are the ones in the Appendices XIV, XV, XVI, XVII which were not considered important in the C.R.P. Analysis.

It would be very unlikely that the answers which were found important in the C.R.P. study and those found by the du Mas method

would be identical. The C.R.P. were not using the reading scores as the focal point of their study. They were not trying to discover items that had a relationship with reading alone. Also their method was to find items which would discriminate the lowest 27 per cent and the highest 27 per cent from the rest. The du Mas method attempted to find items which characterized each criterion score. The C.R.P. was not expecting to get material so rated mathematically that it could be used for predicting anything. Also as du Mas has pointed out in personal communication, Manifest Structure Analysis sometimes chooses items which would be rejected by item analysis.

The distributions were non-normal in shape with great concentrations at the lowest end of the scale. For example, out of 181 cases 53 had a reading score of 1.0. This meant that these could not read at all. The C.R.P. statistician felt that the writer could get better correlation and more effective score prediction if she would eliminate these non-readers from the study. The same opinion were expressed by another statistician from Roosevelt University who examined the writer's unexpected Chi square results. The du Mas method did discriminate among these non-readers and the S score predicted for them might well be an indication as to which ones would make the quickest progress once they had special reading instruction in an E.M.H. class. This might be a suitable subject for further research--namely, which of these pupils, as the present non-readers, would make the most

progress in E.M.H. classes.

Although du Mas does not suggest any other method than those already described, the writer wished to know if her results could be substantiated by some other method. Being advised by a statistician and University professor to use some of the ideas proposed by Siegel (15) in his "Non Parametric Statistics," the writer selected "The Wilcoxon Matched-Pairs Signed-Ranks Test." Siegel had placed in Chapter 5 (The Case of Two Related Samples) all the methods which he considered most useful. The writer was dealing with two related samples. However, all the methods in this chapter were directed toward proving the samples were different. The writer wished to prove her samples were very much alike and so none of the tests were really satisfactory for her purposes. Of the other tests mentioned by Siegel, the McNemar test for the significance of change was too much like the Chi square tests already given to add anything new. The "Sign test" had a poor kind of provision for ties (of which the writer's data had many) and its power-efficiency declined to 63 per cent for large samples. The Wilcoxon test utilized the information about not only the direction of the difference within pairs but the magnitude of the differences. It also provided a good technique for dealing with ties. When the number goes above 25 pairs, the Wilcoxon test assumes that the sum of the ranks,  $T$ , is normally distributed. This is a false assumption so far as the writer's data is concerned. However, in applying the test, the writer



confirmed the results of the Chi square test, namely that the null hypothesis must be rejected. The figures show well beyond the one per cent level of significance that the distributions in the segmental R and S and the clustery R and S for the 181 cases are not similar. It seems to the writer that more research might uncover still other non parametric tests which might not assume a normal distribution for more than 25 cases.

The contribution which this paper hopes to make is to stimulate its readers to experiment with this new method of scaling. Although the writer does not feel that the theory behind the method is able to be readily understood, nevertheless the method is not beyond the abilities of the average teacher or college junior. The scaling frame which the writer has adapted from the Acme Visible File should bring the ownership of a scaling frame within the financial and space limitations of the average teacher or college student. Although the writer used complex, varied, and extensive data in this study, it is hoped that a teacher who composes tests for his classes would see in this method a way of selecting items for his test which will enable him to arrive quickly at a fair grade for each of his students. It might be an adjustment teacher or counselor who needs to cut down on the time denoted to giving and marking tests who would see this method as a time-saver for himself, his pupils, and the school. The factors which showed the closest relation to the reading grade were ones which psychologists, sociologists and educators have agreed

are important in any pupil's success in school. There would probably be lack of agreement if these factors were given rank or weights as are shown in Appendix VIII. As would be expected mental age is the highest ranking factor showing the closest correlation with the reading score. The writer has an hypothesis (which she intends to test) that prediction would be possible without using any patterns to eliminate non essential data merely by using the weights for each category in computing an individual score.

To increase the effectiveness and extend the use of Manifest Structure Analysis, the writer proposes the following:

1. That du Mas in collaboration with his graduate students or colleagues should write a workbook so that those who wish to use his method can develop judgment as to what to do in situations not mentioned in the book.
2. That small scaling frames such as the Acme Visible File used by the writer be used by the people with little money for research (college students and school teachers) who would use Manifest Structure Analysis.
3. That classes and workshops in the theory and application of Manifest Structure Analysis be offered.
4. That du Mas, when and if he rewrites his book, consider the reader for whom he is writing and write at the level of this one specific group.

5. That further research be done to discover the limitations of the method.
6. That research be done to discover other methods of checking the validity and reliability of the results.
7. That no changes be made to destroy the simplicity of the method already established but change be limited to clarifying the theory.
8. That du Mas and his students continue their efforts to extend the number of users of the method, to simplify the explanations of the theory, to multiply the published examples of the actual cases when MSA was used, to reduce the size and cost of the scaling frame.

## BIBLIOGRAPHY

1. Bogardus, E.S. "Measuring Social Distance" Journal of Applied Sociology, 9, 299-308.
2. Coombs, C.H. "Mathematical Models in Psychological Scaling" Journal of the American Statistical Association, 46, 480-489, 1951.
3. Coombs, C.H. Chapter II "Theory and Method of Social Measurement" from Festinger, Leon and Katz, Daniel Research Methods in the Behavioral Sciences. New York, Dryden Press, 1953.
4. du Mas, Frank M. Manifest Structure Analysis. Missoula, Montana, University of Montana, Press, 1955.
5. du Mas, Frank M. "Behavioral Scaling of Personality Tests." Psychological Reports, 1956, 2, 195-204. Southern University Press.
6. du Mas, Frank and King, Mac Bride. "A Manifest Structure Analysis of the Otis S-A Test of Mental Ability, Higher Exam: Form B," Journal of Applied Psychology 42 #4 (August 1958), 269-272.
7. Fisher, R.A. Statistical Methods for Research Workers Vol. V of the set Biological Monographs and Manuals, Oliver and Boyd London, 1950.
8. Garrett, Henry E. Statistics in Psychology and Education 4th Edition Longmans, Green and Co., New York, 1955.
9. Guilford, J.P. Fundamental Statistics in Psychology and Education 3rd Edition, McGraw-Hill Book Co., New York, 1956.
10. Guttman, L. "A Basis for Scaling Qualitative Data" American Sociology Review, 9, 139-150, 1955.
11. Hasse, Derek. A Manifest Structure Analysis of Information Files--Master's Thesis, Missoula, Montana, University of Montana, Press, 1957.
12. Krech, D. and Crutchfield, R.S. Theory and Problems of Social Psychology. New York, McGraw-Hill Book Co., Inc. 1948.
13. Lazarsfeld, Paul F. Mathematical Thinking in the Social Sciences, Chapter 7, A Conceptual Introduction to Latent Structure Analysis. Glencoe, Ill., Free Press, 1955.

14. Remmers, H.H. and Silance, E.B. "Generalized Attitude Scales" Journal of Social Psychology, 5, 298-312, 1934.
15. Siegel, Sidney. Nonparametric Statistics for the Behavioral Sciences, McGraw-Hill Book Co., New York 1956.
16. Stevens, S.S. "On the Theory of Scales of Measurement" Science 1946, 103, 677-680.
17. Thorndike, R.L. and Hagen, E. Measurement and Evaluation in Psychology and Education. New York, John Wiley & Sons, Inc., 1955.
18. Webb, Sam C. "A Generalized Scale for Measuring Interest in Science Subjects" Educational and Psychological Measurement, 11, 456-459, 1951.

# APPENDICES

	Page
1. Announcement of Workshop in <u>Applied Pattern Analysis</u> . .	70
2. The Chicago Project - <u>The Cooperative Research Project</u> William Itkin . . . . .	75
3. A Report on a Preliminary Study with Mentally Handi- capped Children William Itkin . . . . .	78
4. Case Record Folder . . . . .	80
5. Cost Sheet for Acme Visible File . . . . .	86
6. Estimate of Cost of Having the du Mas 55 Inch Scaling Frame Built . . . . .	87
7. The Thirteen Groups of the Clustery Catescale . . . . .	89
8. du Mas' Model and Empirical Analogue . . . . .	90
9. Enlarged Photograph of the Scaling Frame with All the Data Mounted on it . . . . .	91
10. Enlarged Photograph of the Segmental Catescale . . . . .	92
11. Enlarged Photograph of the Intensive Catescale (negative) . . . . .	93
12. Enlarged Photograph of the Intensive Catescale (positive) . . . . .	94
13. Enlarged Photograph of the Clustery Catescale . . . . .	95
14. Categories Used in the Segmental Catescale . . . . .	96
15. Categories Used in the Right Intensive Catescale . . . . .	98
16. Categories Used in the Left Intensive Catescale . . . . .	101
17. Categories Used in the Clustery Catescale . . . . .	104
18. Categories with Weights as Ordered in the Various Catescales . . . . .	107

## APPENDIX I

### ANNOUNCING A WORKSHOP IN

### A P P L I E D P A T T E R N A N A L Y S E S

#### DO YOU WANT TO USE PATTERN ANALYSES TO:

- \*Construct ability, aptitude, achievement, interest or personality tests?
- \*Greatly reduce the number of items in-and time to take-present tests?
- \*Very rapidly evaluate test batteries, profiles or psychographs?
- \*Predict a quantitative or qualitative criterion?
- \*Utilize case histories or application forms as measuring instruments?

#### DO YOU WANT TO:

- \*Save time, space, equipment, personnel, money in your analysis of data?
- \*Do more research than you now do on your present budget?
- \*Maintain your research output under a reduced budget?
- \*Use pattern analyses in your day-to-day evaluations of people?

The emphasis in the workshop will be on the practical applications of configural or pattern analyses to the unidimensional profile.

Bring your own data. We will help you analyze it. Let the savings resulting from research you do at the workshop more than pay for the cost of your attendance. Research projects of moderate size should be completed by the end of the workshop. Observe and participate in the analysis of several practical problems simultaneously. Skilled assistants (college students) available for routine tasks and statistical analyses at very reasonable hourly rates. If you are thinking about a project and do not have your data collected, you will be assisted in the design, procedure for analysis, and practical application.

Excellent cool climate, beautiful surroundings, fine recreational facilities for after-work relaxation.

#### FIRST DAY

Morning. Why pattern or configural analysis? Critical appraisal of classical test methods. Review of major contributors to pattern analysis: Guttman's 'scalogram analysis,' Loevinger's 'homogeneous tests,' Lazarsfeld's 'latent structure analysis,' du Mas' 'manifest structure analysis.' Other contributors: Horst, Lubin, Meehl, Cronbach, Gleser, etc.

Afternoon. Manifest structure analysis. Theory and methods. The scaling frame. The segmental scale. The intensive scale. The clustery scale. (A meeting will follow for participants who bring their own data in order to select assistants, obtain research materials, plan their research so that it will be completed by the end of the workshop).



## SECOND DAY

Morning. (a) Analysis of data by workshop director. Stop-watch demonstration of simplicity, easy and speed of manifest structure analysis: three different scales (segmental, intensive, clustery) will be extracted from 10,000 responses (100 individuals' responses to 100 items) in less than 30 minutes. (b) Analysis of data by workshop participants working together as a group.

Afternoon. (c) Each workshop participant working alone will extract from hypothetical data each of the three kinds of scales: segmental, intensive, clustery. Practice in high speed analysis of data. (d) Those participants who bring their own data can work on their own empirical research at this time. All participants can go from one scaling frame to another and watch and assist in the analyses of empirical data for a wide variety of problems.

## THIRD DAY

Morning. Continuation of second day's work. Predication of a quantitative criterion: manifest structure analysis. Prediction of a qualitative criterion: manifest multichotomy analysis. Some empirical results using both general methods.

Afternoon. Discussion of practical applications of manifest structure analysis and manifest multichotomy analysis to: case histories, application forms, aptitude, ability, achievement, interest and personality tests, job analysis, job evaluation, executive evaluation and development, and other personnel problems in business and industry. A detailed step-by-step review of how to conduct this

Kind of research.

#### FOURTH DAY

Morning. Orientation to the problem of evaluating test batteries. Multiple cutoffs and multiple correlation. Review of some methods of profile analysis: the work of Meehl, McQuitty, du Mas, Cronbach and Gleser, etc. The referent or standard profile. How to obtain referent or standard profiles.

Afternoon. The rapid comparison of an individual's profile with several referent profiles. Methods of comparing profile 'elevations.' Methods of comparing profile 'scatters.' Methods of comparing profile 'shapes.' The interpretation of profiles.

#### FIFTH DAY

Morning. (a) Stop-watch demonstration of simplicity, ease and speed of method for comparing an individual's profile with a referent profile in all three aspects. (b) Workshop participants work together as a group to compare two test battery profiles. (c) Each workshop participant works individually in comparing two profiles. Those participants who bring their own empirical data may work on it at this time. Participants can watch and assist in a variety of profile analyses of actual data.

Afternoon. Application of profile analysis to various problems: evaluation and selection of personnel, attitude and morale survey, effectiveness of training or counselling, degree of similarity between jobs in job analysis, relevance to job evaluation, matching of 'work profiles' with 'worker profiles,' product

comparison and product improvement, etc. Open discussion of a variety of problems until end of seminar.

Exact details regarding the workshop will be worked out to fit the needs of a majority of interested individuals or organizations. The information given below is tentative.

Time: Summer, 1958 (5 Days).

Place: Missoula, Montana

Tuition: \$250.

Registration and Inquiries:

Dr. Frank M. du Mas  
Workshop Director  
Department of Psychology  
Montana State University  
Missoula, Montana

## APPENDIX II

### THE CHICAGO PROJECT

William Itkin

#### Cooperative Research Project

The Chicago Board of Education, the Office of Education of the United States Department of Health, Education, and Welfare, and the Department of Public Instruction of the State of Illinois have entered into an agreement to conduct a cooperative research project in the education of mentally handicapped children. The program is being conducted under the title: How Mentally Handicapped Children Learn Under Classroom Conditions. The project is to be conducted for a period of over three years, from March 18, 1957, until July 30, 1960.

The purpose of the research program is to determine what kinds of instruction and what kinds of classroom organization result in the optimal development of mentally handicapped children. Four points of attack on this complex problem have been selected for study: (1) the effectiveness of special class organization; (2) the importance of individual interests in the reading program; (3) the role of phonics in reading instruction; and (4) the effectiveness of the unit method of instruction with mentally handicapped pupils. The effectiveness of methods of instruction and of types of classroom organization will be studied in relation to the non-academic as well as the academic goals of instruction.

The research program will, therefore, have implications for the instructional program as a whole as well as for the instructional program for the educable mentally handicapped.

Deliberate efforts will be made to enlist the interest and creative ability of the entire teaching force in the research program. It is anticipated that from time to time the project staff will meet problems for which there are as yet no ready solutions. Examples of such problems would be the development of instruments to measure progress in terms of the non-academic goal of the curriculum. Workshops and consultations with members of the teaching, administrative, school psychology, and guidance staffs will be arranged in order that the judgment and creative ability of the teaching and administrative personnel in the field might be directed toward the solution of educational problems of practical importance.

The project staff includes Dr. William Itkin, Director, Dr. Miriam Rooney and Dr. Eugene Richards, psychologists, Dr. Berthold Densch, social research teacher and Mrs. Lucille M. Barry, Miss Mary Cummings and Mrs. Willie H. Scarborough, E.M.H. teachers. A statistician with a background in psychology and education is being sought. The program is under the general direction of Dr. Frances A. Mullen, Assistant Superintendent in charge of Special Education. Dr. David Kopel, Director of the Graduate School of Chicago Teachers College, and Dr. Max Englehart, Director, Division of Student Examinations, are consultants. Miss Bernice M. Grannon, Director, Bureau of Mentally Handicapped Children, Dr. William M.

Canning, Director, Bureau of Child Study; Dr. Carl A. Clark, Head of the Department of Psychology, Chicago Teachers College, and Dr. John M. Beck, Head of the Department of Education of Chicago Teachers College, are ex-officio advisors.

## APPENDIX III

Chicago Board of Education  
Cooperative Research Project

### A REPORT ON A PRELIMINARY STUDY WITH MENTALLY HANDICAPPED CHILDREN

This is a report to Chicago school administrators on a preliminary study being conducted by the Cooperative Research Project during the month of May 1957. The main purpose of this study is to determine which factors are most closely related to the achievement and adjustment of mentally handicapped pupils.

Approximately two hundred fifty pupils were selected at random from E.M.H. classes and from the waiting list. A standardized achievement battery is being administered to this sample by the project staff. At the same time, the pupils in the study are being rated on overall classroom adjustment, behavior, motivation and a number of background factors. Answers to questions such as the following are being sought in the present study: Are there certain distinguishable "types" of mentally handicapped children; i.e., the withdrawn, the lethargic, the hyperactive, the aggressive, the well adjusted? If so, how do these different types of E.M.H. children compare in their academic work? Is the academic work of E.M.H. children influenced by their parents' cooperativeness with the school or by the general educational aspiration level of the family? What

is the effect of parental pressure upon the adjustment and the academic work of mentally handicapped children?

The rating instruments being used in this study were devised with the view of eliciting the maximum amount of pertinent data at a minimum expenditure of teacher time.

The administrator's inquiries and suggestions will be welcomed.

Prepared by:

William Itkin  
Director  
Cooperative Research Project



# APPENDIX IV

Chicago Board of Education  
Cooperative Research Project

## CASE RECORD FOLDER

HOLLERITH CARD  
Col. Punch

Child's Name: \_\_\_\_\_  
                     Last                      First                      Middle  
 Parents or Guardian \_\_\_\_\_ Address \_\_\_\_\_

Hollerith Set No. \_\_\_\_\_

- 10 Supplementary Card #1
- 11 Supplementary Card #2
- 12 Supplementary Card #3

Case No. \_\_\_\_\_ 2-5 \_\_\_\_\_

School \_\_\_\_\_ 6-8 \_\_\_\_\_

District No. \_\_\_\_\_ 9-10 \_\_\_\_\_

Grade - Attitude toward placement - Reason for non-placement \_\_\_\_\_ 11 \_\_\_\_\_

11-9 Kg. 12-5 Sec. Ungr. \_\_\_\_\_ 12 \_\_\_\_\_

11-10 1 C. 12-6 Placed in spite of parental objection

11-11 High Sch. Reg. 12-7 Parental attitude toward placement unknown

12-1 B Semester 12-8 Parent accepted placement

12-2 A Semester 12-10 Not placed because of parental objection

12-3 Lower Ungr. 12-11 Not placed because of unavailability of placement

12-4 Adv. Ungr. Div. 12-12 Not placed because placement not considered urgent; others given priority

Sex - Foreign Language - Background - Cultural \_\_\_\_\_ 13 \_\_\_\_\_

13-1 Boy 13-3 Foreign language spoken in the home \_\_\_\_\_  
 \_\_\_\_\_  
 \_\_\_\_\_

Col. Punch

13-2 Girl  
 13-4 Educational or cultural deprivation  
 13-5 Foreign born

14-3 Puerto Rican - Spanish speaking  
 14-4 Mexican - 1st or 2nd generation  
 14-5 Other Spanish American  
 14-6 North and West European  
 14-7 South and East European  
 14-8 Eastern Asiatic  
 14-9 Western Asiatic and North African  
 14-10 American Indian

14PHYSICAL FACTORS

1 Visual defect	8 Poor health	30
2 Hearing defect	history	
3 Defective teeth	9 Severe physical	
4 Ear, nose, throat defects	deprivation in	
5 Poor motor coordination	early years	
6 Crippled or cardiac	10 Speech defect	
7 Mixed or confused laterality	11 Diagnosed psychia-	
	tric disorder	
	12 Visual or auditory	
	defect sufficient	
	for placement	
1 Diagnosis of organic syn-	6 Gross locomotor	31
drome by neurologist or	disturbances	
endocrinologist	after age 4	
2 Head injury at birth	7 Epilepsy, seizures	
3 Deficient animation at	diagnosed	
birth	8 Severe head	
4 Meningitis or encephalitis	injuries	
5 Abnormal neurological		
signs		

ACHIEVEMENTS

Oral reading score	12 Recent BCS school test data	32-33
Reading comprehension	34-11 Primary; 34-12 Elementary	34-35
grade score	35-12 Recent BCS or school data	
Reading vocabulary test	36-11 Primary; 36-12 Elementary	36-37
grade score	37-12 Recent BCS or school data	
Average reading score		38-39

ACHIEVEMENTSCol. Punch

Arithmetic reasoning test 40-11 Primary; 40-12  
grade score Elementary  
41-12 Recent BCS or school  
data

40-41 \_\_\_\_\_

Arithmetic computation test  
grade score 42-11 Primary; 42-12  
Elementary  
43-12 Recent BCS or school  
data

42-43 \_\_\_\_\_SCHOOL HISTORY44 \_\_\_\_\_

Number of schools attended 10 Previously in  
6 or more, punch 6 physically handi-  
capped classroom

7 Previous private day school  
placement 11 Formerly in EMH  
division; now in  
8 Residential school history regular grades

9 Previously in social 12 Previously excused  
adjustment center from school

Age at EMH placement (Punch 1 and 3 for 13, 1 and 4 for  
14, etc.)

45 \_\_\_\_\_

Home School

46-48 \_\_\_\_\_

Transiency index of home school

49 \_\_\_\_\_

Speech Appraisal Score

50 \_\_\_\_\_

School pupil attended 1B

Code 12 for Speech Problem indicated in 1B appraisal

Punch Col.

15-17 Date of Birth: \_\_\_\_\_ (Punch year and month)  
Year Month Day

18 Date of most recent B.C.S. Examination  
(Year)

19 Test  
1 Rev St. -B L 3 W.-B  
2 Rev St. -B M 4 WISC

punch      Col.

## Test

## 5 Ontario

6 Kuhlman Ment. Development

## 7 Stanford Binet

## Sight Saving

9 Others

20-21. C A as of \_\_\_\_\_

22-23 M A as of \_\_\_\_\_ (computed to date)

24-25 I Q

25-11 I Q above 100

25-12 I Q questioned

26 Marital Status of Parents

1 Living with both parents

2 Temporary separation (11) living with mother  
(12) with father

3 Permanent separation (11) living with mother  
(12) with father

4 Divorced (11) living with mother  
(12) with father

5 (11, 12) Living with one parent; other deceased

6 Living with one parent and stepparent (also  
code 2,3,4 or 5 whichever applies)

7 Living with blood relatives

Living with adoptive parent or parents since infancy (code other numbers which apply)

10 Living with adoptive parents; adopted at age 3 or later (code other numbers which apply)

27 Living with foster parents or in a children's home

Code Nos. 1 to 7 incl. and 11 and 12 which apply  
as in Col. 27

27-9 Living in Children's Home.-Own parents living,  
known to child and visiting.

27-10 Living in Children's Home - Own parents deceased, not known to child, or whereabouts unknown 2 years or more.

Punch    Col.

23      Known to social agencies

- |   |                              |   |                   |
|---|------------------------------|---|-------------------|
| 1 | ADC                          | 7 | IJR               |
| 2 | Other public relief agencies | 8 | Hospital or other |
| 3 | Private relief agencies      |   | child guidance    |
| 4 | Receiving pension            |   | clinics           |
| 5 | Family court                 | 9 | Private psychia-  |
| 6 | Family service bureaus       |   | tric arrangement  |

29      Chicago residence

Number of years child has lived in the city  
 11 (11 or more)  
 12 All his life

#### ADJUSTMENT RATING SCALE

51      Overall Classroom Work

52      Overall Classroom Behavior

53      Adjustment to other Children

54      Social Participation

55-56    Total Adjustment Rating Scale

57-58    Total Trait Rating Scale

#### SOCIAL FACTORS

59      S-2 Status Rating

60      Sims Adaptation Score

61      Parental Cooperativeness Rating

62      Home Atmosphere Rating

63      Educational Aspiration Rating

64      Community Rating

#### PERSONALITY TESTS

65      Central Personality Pattern

66-67    Children's Manifest Anxiety Scale

<u>Punch</u>	<u>Col.</u>	<u>PERSONALITY TESTS</u>
<u>      </u>	<u>68</u>	Pressure for Achievement
<u>      </u>	<u>69-70</u>	Total Motivation Score
<u>      </u>	<u>71</u>	Haggerty-Olson Wickman Behavior Rating Schedule B (Refer to class Interval Chart)

EMOTIONAL MATURITY SCALE

<u>      </u>	<u>72</u>	
<u>      </u>	<u>73</u>	
<u>      </u>	<u>74</u>	
<u>      </u>	<u>75</u>	
<u>      </u>	<u>76</u>	
<u>      </u>	<u>77</u>	
<u>      </u>	<u>78</u>	
<u>      </u>	<u>79-80</u>	<u>SOCIAL MATURITY SCORE</u>

## APPENDIX V

### COST OF THE ACME VISIBLE FILE

#### A SUBSTITUTE FOR THE SCALING FRAME

1.	Large Size Light Aluminum Two Leaf or Fold Over File with 21 by 14 Inch Leaf Size.	
	List Price with 98 Pockets. . . . .	\$40.10
2.	Extra Pockets . . . . .100 for. . .	29.25
3.	Perforated Title Inserts. . . . .per 100. . .	1.45
4.	10T Flags (if used) . . . . .per 1000 . . 12 1/2% Discount Allowed if Bought in Quantity)	12.95
5.	Marking Ink (if flags are not used) . . . . .	.50
	Price Range for Scaling 100 Cases (not using flags) .	71.30
	If Flags are Used . . . . .up to . . .	138.80

## APPENDIX VI

### ESTIMATE OF COST OF BUILDING THE DU MAS

#### 55 INCH SCALING FRAME

Informant: Estimation Department of Geo. Sollett Co. Contractors

#### A. COST OF MATERIALS

1. Cost of plywood 3/4 inch thick  
25 sq. feet needed but 2 pieces 8 ft. by 4 ft.  
or 64 sq. feet must be bought at 40¢ per  
square foot . . . . . \$25.60
2. Cost of masonite 1/4 inch thick 25 square feet  
needed but 2 (3 ft. by 4 ft.) pieces must be  
purchased at 20¢ per square foot. . . . . 12.80
3. Hinges, hooks, and other hardware . . . . . 2.00  
\$40.40

#### B. LABOR COSTS

1. If Union Carpenters do the Work
  - a. estimate of time for building  
board . . . . . 16 hours
  - b. estimate of time for drilling  
holes at 20 per hour. . . . 500 hours  
Total time for carpenter to  
do the work is . . . . . 516 hours
  - c. union carpenters wages \$3.65 per hour plus  
\$1.40
    1. insurance . . . . . 15 per cent
    2. overhead . . . . . 10 per cent
    3. profit . . . . . 10 per cent
    4. tools & electricity. 5 per cent  
40% or \$5.05

Total \$5.05 X 516 hours. . . . . \$2605.80



2. If the Writer's Husband's Employees Who Make  
Electronic Equipment do the Work

- a. estimate of time for building  
board . . . . . 16 hours
- b. estimate of time for drilling  
holes 20 per minute . . . . . 8 1/3 hours
- c. total time 24 1/3 hours at \$4.00  
per hour . . . . . \$97.33

C. TOTAL COST WOULD RANGE FROM \$137.73 to \$2,646.20

# APPENDIX VII

C 1	C 2	C 3	C 4	C 5	C 6	C 7	C 8	C 9	C 10	C 11	C 12	C 13
5	5	5	5	5	5	5	5	5	5	5	5	5
7	7	7	7	7	7	7	7	7	7	7	7	7
11	11	11	11	11	11	11	11	11	11	11	11	11
12	12	12	12	12	12	12	12	12	12	12	12	12
13	13	13	13	13	13	13	13	13	13	13	13	13
17	17	17	17	17	17	17	17	17	17	17	17	17
18	18	18	18	18	18	18	18	18	18	18	18	18
22	22	22	22	22	22	22	22	22	22	22	22	22
25	25	25	25	25	25	25	25	25	25	25	25	25
28	28	28	28	28	28	28	28	28	28	28	28	28
32	32	32	32	32	32	32	32	32	32	32	32	32
33	33	33	33	33	33	33	33	33	33	33	33	33
38	38	38	38	38	38	38	38	38	38	38	38	38
41	41	41	41	41	41	41	41	41	41	41	41	41
44	44	44	44	44	44	44	44	44	44	44	44	44
53	53	53	53	53	53	53	53	53	53	53	53	53
54	54	54	54	54	54	54	54	54	54	54	54	54
60	60	60	60	60	60	60	60	60	60	60	60	60
62	62	62	62	62	62	62	62	62	62	62	62	62
63	63	63	63	63	63	63	63	63	63	63	63	63
80	80	80	80	80	80	80	80	80	80	80	80	80
83	83	83	83	83	83	83	83	83	83	83	83	83
84	84	84	84	84	84	84	84	84	84	84	84	84
85	85	85	85	85	85	85	85	85	85	85	85	85
87	87	87	87	87	87	87	87	87	87	87	87	87
90	90	90	90	90	90	90	90	90	90	90	90	90
91	91	91	91	91	91	91	91	91	91	91	91	91
96	96	96	96	96	96	96	96	96	96	96	96	96
99	99	99	99	99	99	99	99	99	99	99	99	99
118	118	118	118	118	118	118	118	118	118	118	118	118
120	120	120	120	120	120	120	120	120	120	120	120	120
123	123	123	123	123	123	123	123	123	123	123	123	123
125	125	125	125	125	125	125	125	125	125	125	125	125
129	129	129	129	129	129	129	129	129	129	129	129	129
132	132	132	132	132	132	132	132	132	132	132	132	132
134	134	134	134	134	134	134	134	134	134	134	134	134
143	143	143	143	143	143	143	143	143	143	143	143	143
148	148	148	148	148	148	148	148	148	148	148	148	148
149	149	149	149	149	149	149	149	149	149	149	149	149
150	150	150	150	150	150	150	150	150	150	150	150	150
151	151	151	151	151	151	151	151	151	151	151	151	151

# APPENDIX VIII

## THE MODEL

The Normal Distribution

Properties

$\alpha$  . . . . .

$\beta$  . . . . .

$\gamma$  . . . . .

$\delta$  . . . . .

.

.

.

## THE EMPIRICAL ANALOGUE

The Empirical Distribution

Properties

. . . . . A

. . . . . B

. . . . . C

. . . . . D

.

.

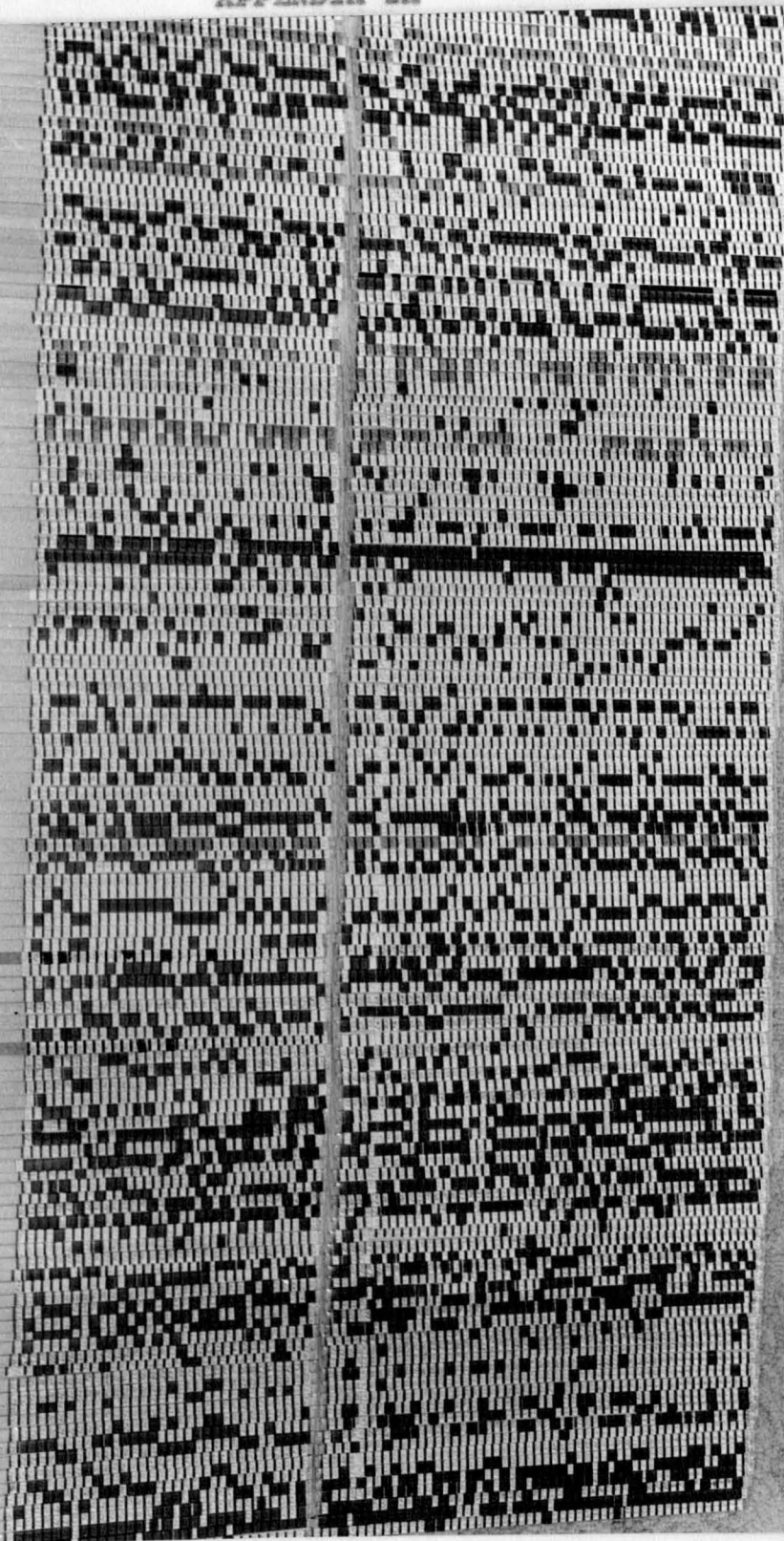
.

Collating Operations:  $\chi^2$

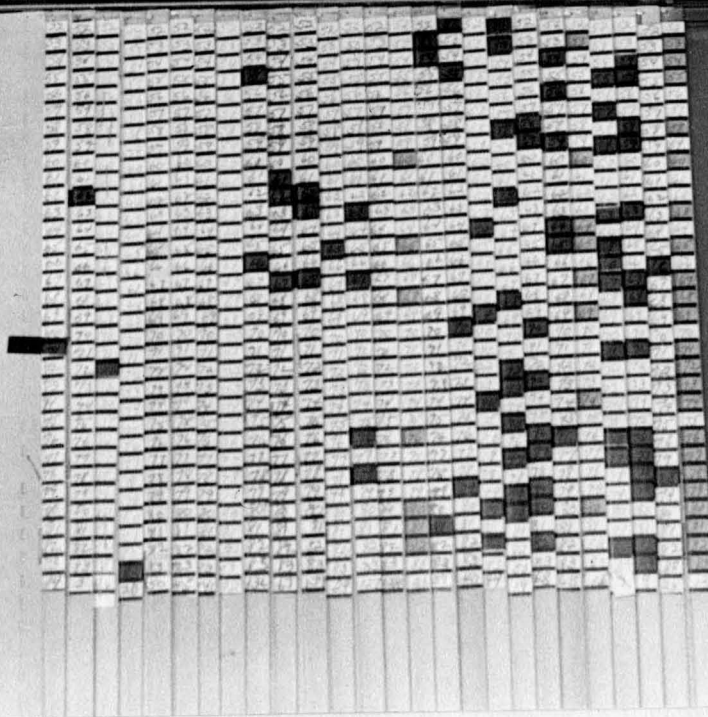
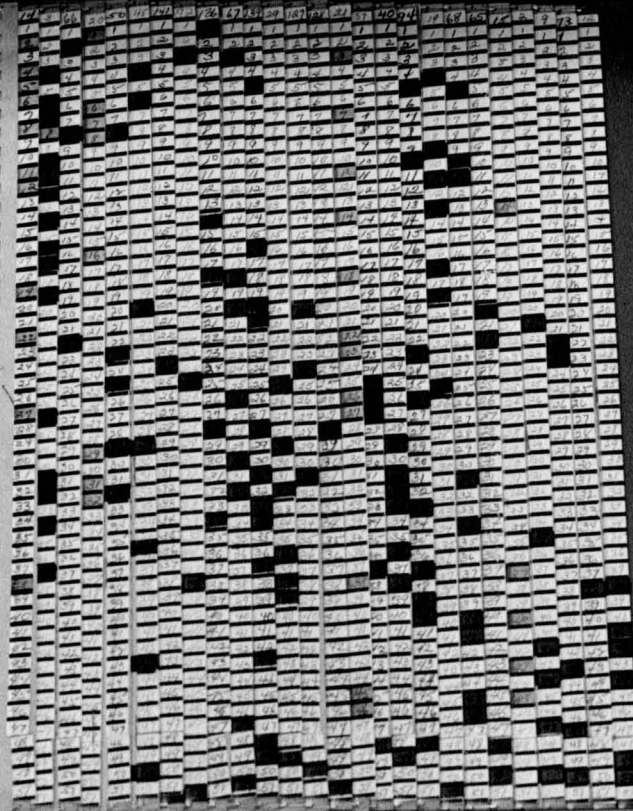
Criterion:  $P_{\chi^2} = .05$

An example of the application of a model to empirical data.

APPENDIX IX

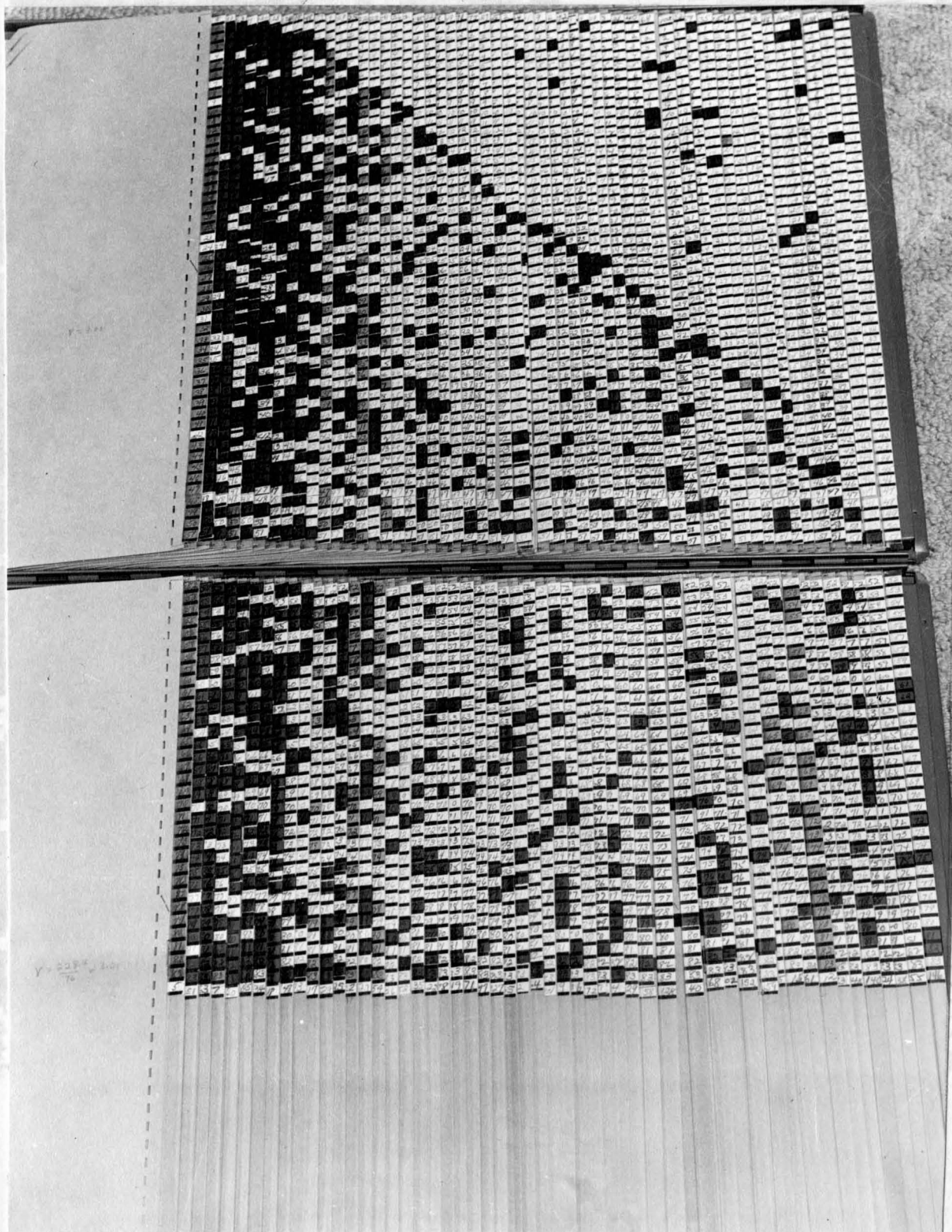


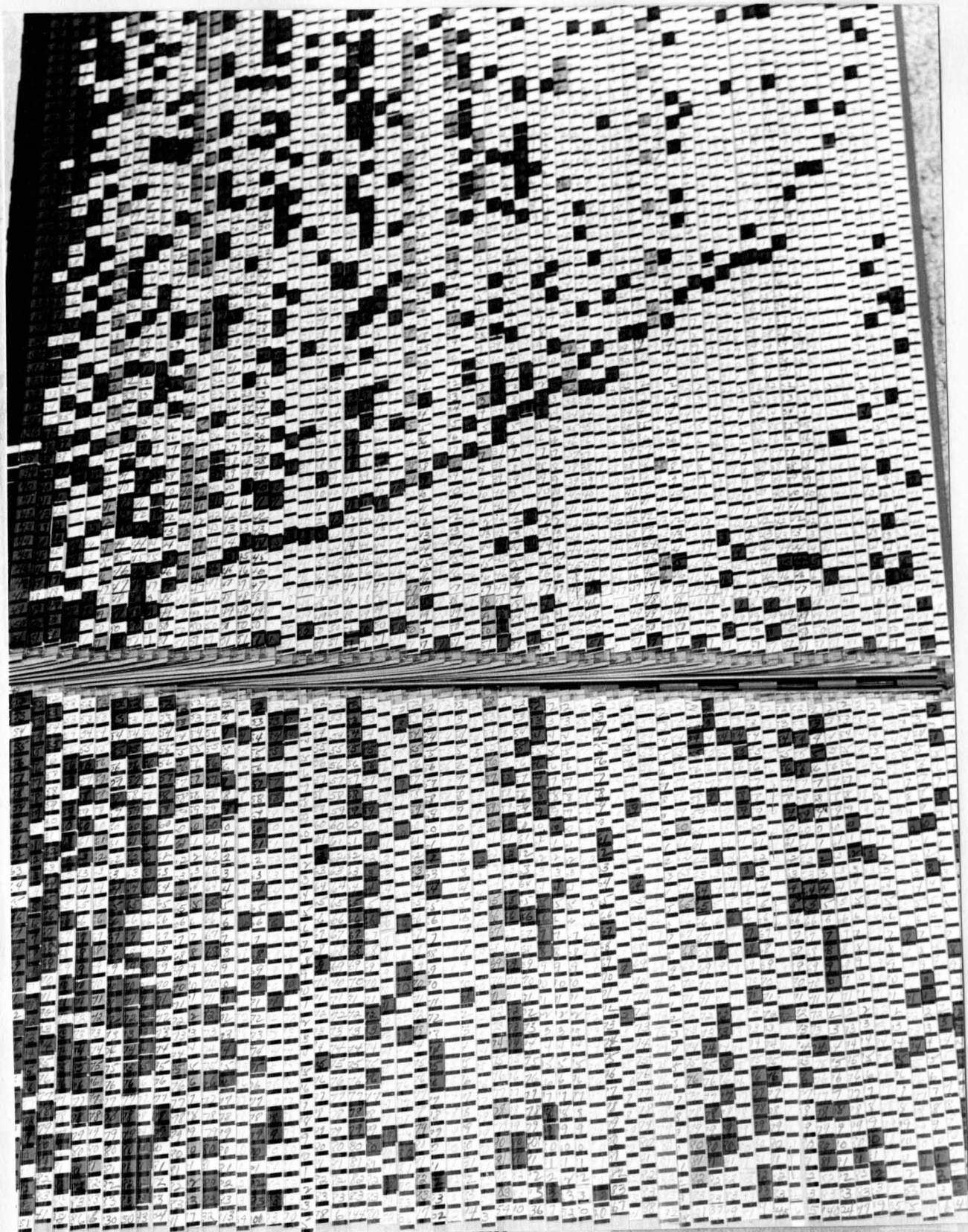
APPENDIX X





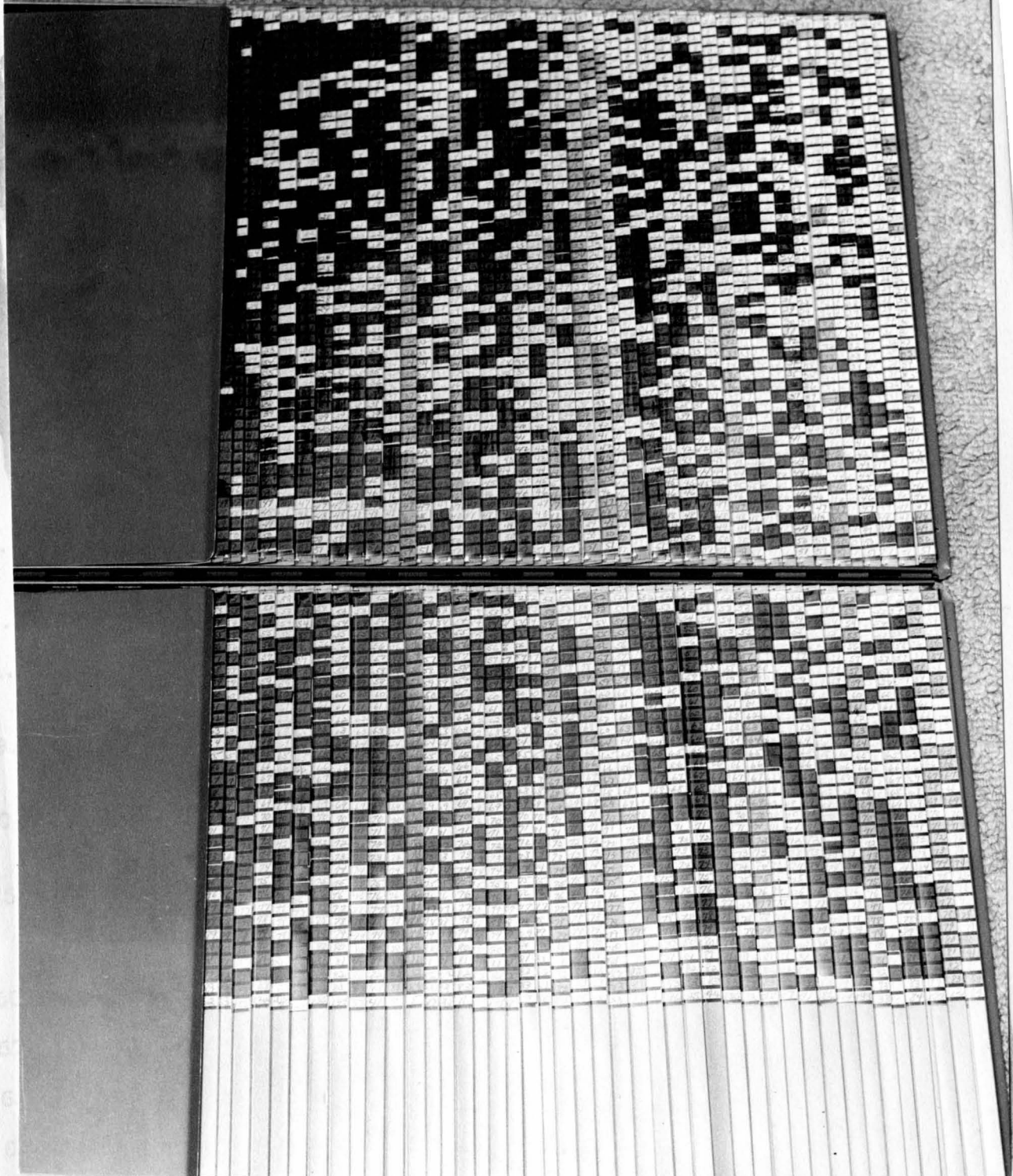
APPENDIX XI







APPENDIX XIII



67. Item 49 schools in transition class 1. P. 95  
equally good neighborhoods as the above)



## APPENDIX XIV

### CATEGORIES USED IN THE SEGMENTAL CATESCALE

2. Item 12-3 Lower ungraded.
3. Items 12-1 and 12-2 Grades 6B - 8A.
9. Items 20 and 21 Chronological ages eight years and 5 months to eleven years 12 months.
12. Items 22 and 23 Mental ages 5 years and ten months to seven years 12 months.
14. Items 22 and 23 Mental ages 10 years and 0 months to thirteen years 12 months.
19. Item 12-11 Not placed because of unavailability of placement.
- \*20. Item 12-12 Placement not urgent.
21. Socio-economic Rating Scale I (economic status of the family) - poverty stricken or on relief.
- \*29. Item 13-3 Foreign born or foreign language spoken in the home.
40. Item 63 or Socio-economic Rating Scale I (education) not considered important.
45. Socio-economic Rating Scale V (cultural level of the family) (3) belong to lodges; play golf on public course; regular bowling night, etc.
- \*50. Item 28-3 Private relief agencies.
- \*57. Items 30-5 or 6 Crippled or poor motor coordination.
65. Item 44-12 Previously excused from school.
66. Item 49 Schools in transiency class I. (Small schools in fine neighborhoods).
67. Item 49 Schools in transiency class 2. (Larger schools in equally good neighborhoods as the above).

- 68. Item 49 Schools in transiency class 3. (Schools not as well placed as the above but still good schools).
- 73. Item 49 Schools in transiency class 8 to 10. The schools with highest transiency records and least desirable socio-economic status.
- 92. Items 55 and 56 Total adjustment rating 16 and above.
- 94. Item 76 Emotional Maturity III scores 31-43.
- 127. Adaptation of Sims scorecard 16B. Mother does skilled work - usually in a factory.
- 128. Adaptation of Sims scorecard 16B. Mother does managerial work.
- 136. Item 64 Community rating 6 and 7 slightly below the middle of the community ratings.
- \*139. Adaptation of Sims scorecard 19. Father went to high school.
- \*141. Adaptation of Sims scorecard 18. Mother went to college.

## APPENDIX XV

### CATEGORIES USED IN THE RIGHT INTENSIVE CATESCALE

2. Item 12-0 Lower ungraded.
4. Item 12-1 and 12-2 Grades 1A to 5A.
5. Item 19-1 The Revised Stanford-Binet.
7. Item 65 Central Personality Patterns (scores 1, 2, 3, 4).
9. Item 20-21 Chronological ages 8 years 5 months to 11 years 12 months.
12. Items 22 and 23 Mental ages 5 years and 10 months to 7 years 12 months.
15. Item 24-25 I.Q. 50-59.
17. Item 24-25 I.Q. 70-78.
22. Socio-economic Rating Scale I (economic status of family) - poor.
25. Socio-economic Rating Scale \*II (occupation of family head) - unskilled, low paying but steady.
27. Item 13-2 Girl.
- \*29. Item 13-3 Foreign born or foreign language spoken in the home.
- \*35. Item 27 Lives with foster parents or in an orphanage.
36. Socio-economic Rating Scale III. (Residence of family) - slums or public housing project.
40. Item 63 or Socio-economic Rating Scale I (education) - not considered important.
- \*40. Item 28-2 Other public relief agencies.
- \*52. Item 28-8 or 9 Hospital or other child guidance clinics or private psychiatric arrangement.

- \*54. Item 30-1 Visual defects.
- \*55. Item 30-2 Hearing defects.
- \*57. Items 30-5 or 6 Crippled or poor motor coordination.
- \*58. Item 30-8 Poor health history or physical deprivation.
- \*59. Item 30-10 Speech defects.
- \*61. Items 31-1-3 Very serious brain damage or nerve damage.
- 63. Item 44 Attended only 1 to 5 different schools.
- 64. Item 44 Number of schools attended 6-10.
- 65. Item 44-12 Only the present school - previously excused.
- 68. Item 49 Schools in transiency class 3 (see Item 49 in Segmental Scale List - Appendix XIV).
- 71. Item 49 Schools in transiency class 6.
- 72. Item 49 Transiency classification 7 (below the middle but not the very poor).
- 73. Item 49 Schools in transiency class 8 to 10. The schools with highest transiency records and least desirable socio-economic status.
- 77. Item 51 Score for classroom work=0.
- 78. Item 51 Overall classroom work score 1 - not too good but better than zero rating.
- 93. Item 79-80 Emotional Maturity IV scores 10-20.
- 101. Item 60 Sims Adaption Score. (Born in northern states).
- 102. Item 71 Haggerty-Olson Wichman Behavior Rating Schedule B scores 3 or 4.
- 118. Item 59 Socio-economic status rating 1-4.
- 120. Item 60 Sims Adaption Score 1-5.
- 122. Item 61 Parental Cooperativeness Rating 1 and 2.
- 124. Item 61 Parental Cooperativeness Rating 5 and 12.

- 125. Adaptation of Sims scorecard 16B. Mother does not work outside the home.
- 126. Adaptation of Sims scorecard 16B. Mother works at service jobs (hospitals, housework, waitress, saleswoman).
- 127. Adaptation of Sims scorecard 16B. Mother does skilled work - usually in a factory.
- 128. Adaptation of Sims scorecard 16B. Mother does managerial work.
- 129. Item 62 Home atmosphere rating 1 and 2.
- 134. Item 64 Community Rating 1 and 2.
- 135. Item 64 Community Rating 3 and 4.
- \*138. Adaptation of Sims scorecard 17. Father went to college.
- \*140. Adaptation of Sims scorecard 18 and 20. Mother's education unknown.
- 145. Item 68 Pressure for achievement scores 5 and 6 (high).
- 146. Item 69-70 Total motivation scores 1-5 (low).

## APPENDIX XVI

### CATEGORIES USED IN THE LEFT INTENSIVE CATESCALE

3. Items 12-1 and 12-2 Grades 6B - 8A.
4. Items 12-1 and 12-2 Grades 1A - 5A.
- \*3. Items 19-3, 4, 5, 6, 9 All other forms of intelligence tests than the Stanford-Binet.
9. Item 20-21 Chronological ages 3 years 5 months to 11 years 12 months.
10. Item 20-21 Chronological ages 12 years 0 months to 13 years 12 months.
11. Item 20-21 Chronological ages 14 years 0 months to 16 years 3 months.
14. Item 22-23 Mental ages 10 years 0 months to 13 years 12 months.
16. Item 24-25 I.Q. 60-69.
17. Item 24-25 I.Q. 70-78.
- \*19. Item 12-11 Not placed because of unavailability of placement.
- \*20. Item 12-12 Placement not urgent.
21. Socio-economic Rating Scale I - (education) not considered important.
28. Item 13-1 Boy.
- \*32. Item 26-11 Lives with mother.
- \*34. Item 26-12 Lives with father or other relative.
- \*35. Item 27 Lives with foster parents or in an orphanage.
36. Socio-economic Rating Scale III - Residence of family 1 - slums or public housing project.
41. Socio-economic Rating Scale IV - Education aspirations of the family 3 and 4 hope for high school or college.

45. Socio-economic Rating Scale V - (Cultural level of the family) 2- belong to lodges, play golf on public golf courses, have regular bowling nights, etc.
- \*47. Item 26-3 or 26-4 Parents divorced or have permanent separation.
- \*49. Item 28-2 Other public relief agencies.
- \*50. Item 28-3 Private relief agencies.
- \*51. Item 28-5 or 7 Institute of Juvenile Research or Family Court.
- \*52. Item 28-8 or 9 Hospital or other child guidance clinics or private psychiatric arrangement.
- \*54. Item 30-1 Visual defects.
- \*55. Item 30-2 Hearing defects.
- \*59. Item 30-10 Speech defects.
62. Item 30 None of these defects.
64. Item 44 Number of schools attended 6-10.
65. Item 44-12 Only the present school - previously excused.
66. Item 49 Transiency classification of school 1 (small schools in exclusive neighborhoods).
67. Item 49 Transiency classification 2 (larger schools in exclusive neighborhoods).
70. Item 49 Transiency classification 5 (middle - neither very poor nor very fine).
72. Item 49 Transiency classification 7 (below the middle but not the very poor).
78. Item 51 Overall classroom work score 1 - not too good but better than zero rating.
80. Item 51 Overall class work score 3 - middle in achievement.
86. Item 54 Social participation score 1 and 2 (low).
88. Item 54 Social participation scores 5 and 6 (high).

- 92. Item 55 and 56 Total adjustment rating scale scores 16 and up (highest).
- 93. Item 79-80 Emotional Maturity IV scores 10-20.
- 100. Adaptation of Sims scorecard 15D. Born and raised in southern states.
- \*102. Item 71 Haggerty-Olson Wichman Behavior Rating Schedule B scores 3 or 4.
- \*104. Item 71 Haggerty-Olson Wichman Behavior Rating Schedule B scores 7 and 8.
- 113. Socio-economic Rating Scale IV - Educational aspirations of family 12 - unknown.
- 117. Item 77 Emotional Maturity III score 11-20.
- 124. Item 61 Parental Cooperativeness Rating.
- 126. Adaptation of Sims scorecard 16B. Mother works at service jobs (hospitals, housework, waitress, saleswoman).
- 128. Adaptation of scorecard 16B. Mother does managerial work.
- 130. Item 62 Home atmosphere rating - scores 3-4 (middle).
- 136. Item 64 Community rating scores 6 and 7.
- \*137. Adaptation of Sims scorecard 17 and 19. Father's educational level unknown.
- \*140. Adaptation of Sims scorecard 18 and 20. Mother's education unknown.
- 143. Item 68 Pressure for achievement scores 1 and 2 (low).
- 144. Item 68 Pressure for achievement scores 3 and 4 (medium).
- 145. Item 68 Pressure for achievement scores 5 and 6 (high).
- 146. Item 69-70 Total motivation scores 1-5 (low).
- 150. Item 73 Emotional Maturity Scale I - scores 21-32 (high).
- 151. Item 74 Emotional Maturity Scale II - scores 10-20.
- 152. Item 75 Emotional Maturity Scale II - scores 21-30.



## APPENDIX XVII

### CATEGORIES USED IN THE CLUSTERY PATTERNS

- 5. Item 19-1 Revised Stanford-Binet Form L.
- 7. Item 65 Central Personality Patterns (scores 1, 2, 3, 4).
- 11. Item 20-21 Chronological ages 14 years 0 months to 16 years 3 months.
- 12. Item 22-23 Mental age 5 years 10 months to 7 years 12 months.
- 13. Item 23-24 Mental age 8 years 0 months to 9 years 12 months.
- 17. Item 24-25 I.Q. 70 to 88.
- 18. Item 12-18 Parents accepted placement in ungraded divisions.
- 22. Socio-economic Rating Scale I - (economic status of family) - poor.
- 25. Socio-economic Rating Scale \*II - (occupation of family head) - listed as unskilled or low paying but steady.
- 28. Item 13-1 Boy.
- \*32. Item 26-11 Lives with mother.
- \*33. Item 26-1 Lives with both parents.
- 38. Socio-economic Rating Scale III - Residence of the family lives in rented homes or apartments not too far from the slums or may own or rent in respectable parts of the city.
- 41. Socio-economic Rating Scale IV - Educational aspirations of the family 3 and 4 high school graduation hoped for or taken for granted.
- 44. Socio-economic Rating Scale V - Cultural level of the family characterized by frequent trips to free museums, zoos, parks, daytime movies, etc.
- \*53. Item 28 Known to none of the social agencies mentioned in 28 (1-9).

- \*54. Item 30-1 Visual defects.
- 60. Item 30 No physical defects mentioned in 30 (1-12).
- 62. Item 31 No physical defects mentioned in 31 (1-8).
- 63. Item 44 Attended only 1 to 5 different schools.
- 80. Item 51 Had a score of 3 for classroom work.
- 82. Item 52 Had a score of 3 or 4 for classroom behavior.
- 85. Item 53 Had a score of 3 or 4 for adjustment to other children.
- 87. Item 54 Had a score of 3 or 4 for social participation.
- 90. Item 55-56 Had a score of 6 to 10 on the total adjustment rating scale.
- 96. Item 29 Has lived whole life or all of school life or 11 to 14 years in Chicago.
- 99. Socio-economic Rating Scale 15B. Born in Chicago.
- 118. Item 59 Total or summary of socio-economic status characterized by ratings 1, 2, 3, or 4.
- 120. Item 60 Sims adaption score of 1, 2, 3, 4, or 5.
- 123. Item 61 Parental Cooperativeness Ratings 3 (average) and 4 (more cooperative than average).
- 125. Adaptation of Sims scorecard Item 16A. Mother does not work outside of the home.
- 129. Item 62 Home atmosphere ratings 1 poor or 2, fair.
- 132. Item 63 Educational aspirations 1 education not considered important, 2 children not expected to go to high school, or 3 high school graduation hoped for.
- 134. Item 64 Community rating 1 or 2 - live in slums or in public housing projects.
- 143. Item 68 Pressure for achievement 1 or 2 either hostility toward school attendance or indifference.

148. Item 68 to 70 Total motivation scores 11-16.
149. Item 72-73 Emotional Maturity scores 9-20.
150. Item 74-75 Emotional Maturity scores 21-32.
151. Item 76-77 Emotional Maturity scores 10-20
- 1st pattern
- 2nd pattern

# APPENDIX XVIII

## CATESCALES

### CATEGORIES WITH WEIGHTS ORDERED BY POSITION IN PATTERN

#### SEGMENTAL

14 = 4.38  
 3 = 4.22  
 66 = 4.12  
 20 = 4.11  
 50 = 4.02  
 45 = 3.90  
 141 = 3.90  
 92 = 3.80  
 136 = 3.74  
 67 = 3.30  
 139 = 3.25  
 29 = 3.21  
 127 = 3.20  
 128 = 3.20  
 21 = 3.00  
 57 = 2.96  
 40 = 2.88  
 94 = 2.83  
 19 = 2.80  
 68 = 2.65  
 65 = 2.60  
 15 = 2.45  
 2 = 2.25  
 9 = 2.13  
 73 = 2.13  
 12 = 2.06

#### INTENSIVE (RT)

5 = 3.01	64 = 2.91
151 = 3.14	101 = 2.92
63 = 3.17	9 = 2.13
7 = 3.03	36 = 2.93
120 = 3.15	72 = 3.80
125 = 3.18	57 = 2.96
134 = 3.26	4 = 2.47
110 = 3.03	29 = 3.21
148 = 3.41	58 = 2.11
93 = 3.39	128 = 3.20
17 = 3.44	40 = 2.88
129 = 3.05	68 = 2.64
25 = 3.14	102 = 2.30
22 = 3.01	152 = 3.20
27 = 2.89	59 = 2.91
54 = 2.71	35 = 3.70
126 = 2.96	15 = 2.45
52 = 2.92	61 = 2.10
135 = 2.78	12 = 2.06
122 = 2.74	73 = 2.13
78 = 2.83	145 = 2.80
49 = 2.30	140 = 2.45
71 = 3.10	124 = 2.92
77 = 2.77	138 = 2.70
127 = 3.20	55 = 3.70
65 = 2.60	146 = 1.55
2 = 2.25	

# CATESCALES

## CATEGORIES WITH WEIGHTS ORDERED BY POSITION IN PATTERN

### INTENSIVE (LEFT)

62 = 3.23	10 = 2.92
151 = 3.14	136 = 3.74
41 = 3.37	117 = 3.13
28 = 3.38	152 = 3.29
86 = 2.99	50 = 4.02
16 = 3.94	20 = 4.11
130 = 3.60	67 = 3.30
150 = 2.61	101 = 2.92
143 = 2.78	88 = 3.66
104 = 3.31	72 = 3.80
11 = 3.75	128 = 3.20
47 = 3.40	21 = 3.00
32 = 3.35	137 = 2.36
113 = 2.73	19 = 3.80
34 = 3.50	51 = 3.30
100 = 3.17	9 = 2.13
93 = 3.39	145 = 2.80
70 = 3.55	66 = 4.12
45 = 3.90	65 = 2.60
73 = 2.83	140 = 2.45
6 = 2.81	124 = 2.92
144 = 3.42	49 = 2.30
80 = 3.44	35 = 3.70
52 = 2.92	55 = 3.70
126 = 2.96	59 = 2.91
17 = 3.44	36 = 2.93
102 = 2.30	64 = 2.91
92 = 3.80	4 = 2.47
14 = 4.38	146 = 1.55
3 = 4.20	
54 = 2.71	

### CLUSTERY (1) LOWEST

17 = 3.44
23 = 3.38
32 = 3.35
134 = 3.26
62 = 3.23
125 = 3.18
63 = 3.17
120 = 3.15
151 = 3.14
113 = 3.09
129 = 3.05
7 = 3.03
18 = 3.03
84 = 3.02
5 = 3.01
90 = 3.01
96 = 3.01
143 = 2.78
150 = 2.61
12 = 2.06

### CLUSTERY (13) HIGHEST

149 = 3.77
11 = 3.75
53 = 3.56
60 = 3.46
80 = 3.44
17 = 3.44
38 = 3.41
23 = 3.38
41 = 3.37
32 = 3.35
123 = 3.33
134 = 3.26
83 = 3.18
85 = 3.18
125 = 3.18
132 = 3.18
91 = 3.15
120 = 3.15
87 = 3.14
151 = 3.14
7 = 3.03
33 = 2.76
150 = 2.61

APPROVAL SHEET

The dissertation submitted by Lucille Armstrong Foster has been read and approved by five members of the Department of Education.

The final copies have been examined by the director of the dissertation and the signature which appears below verifies the fact that any necessary changes have been incorporated, and that the dissertation is now given final approval with reference to content, form, and mechanical accuracy.

The dissertation is therefore accepted in partial fulfillment of the requirements for the degree of Doctor of Education.

December 31, 1958  
Date

Samuel T. Mayo  
Signature of Adviser